

A resource theory of superposition

T. Theurer,¹ N. Killoran,^{1,2} D. Egloff,¹ and M.B. Plenio¹

¹*Institut für Theoretische Physik, Albert-Einstein-Allee 11, Universität Ulm, 89069 Ulm, Germany*

²*Department of Electrical and Computer Engineering,
University of Toronto, Toronto, ON M5S 3G4, Canada*

(Dated: April 3, 2017)

The superposition principle lies at the heart of many non-classical properties of quantum mechanics. Motivated by this, we introduce a rigorous resource theory framework for the quantification of superposition of a finite number of linear independent states. This theory is a generalization of resource theories of coherence. We determine the general structure of operations which do not create superposition, find an elementary connection to unambiguous state discrimination, and propose several general quantitative superposition measures. We show that several main results from resource theories of coherence still hold in our more general setting. Of special importance are two results about the free completion of trace decreasing operations and the free probabilistic transformation between pure states that are also valid for the special case of coherence.

Introduction. – During the last decades, there has been an increasing interest in quantum technologies. The main reason for this are the operational advantages of protocols or devices working in the quantum regime over those relying on classical physics. Early examples include quantum cryptography [1], quantum dense coding [2] and quantum teleportation [3], where entanglement acts as a resource which is consumed and manipulated. Therefore the detection, manipulation and quantification of entanglement was investigated, leading to the resource theory of entanglement [4]. Typical quantum resource theories (QRTs) are built by imposing an additional restriction to the laws of quantum mechanics [5–7]. In the case of entanglement theory, this is the restriction to local operations and classical communication (LOCC). From such a restriction, the two main ingredients of QRTs emerge: The free operations and the free states (which are LOCC and separable states in the case of entanglement theory). All states which are not free contain the resource under investigation and are considered costly. Therefore free operations must transform free states to free states, allowing for the resource to be manipulated but not freely created. Once these main ingredients are defined, a resource theory investigates the manipulation, detection, quantification and usage of the resource.

In principle, not only entanglement but every property of quantum mechanics not present in classical physics could lead to an operational advantage [8, 9]. This motivates the considerable interest in quantifying non-classicality in a general way [10–14]. The superposition principle underlies many non-classical properties of quantum mechanics including entanglement or coherence. Recently resource theories of coherence [11, 15, 16] attracted a lot of attention. In these settings, the free states form a finite orthonormal basis of the system under consideration and the resource is the superposition of these, called coherence. Here we present a generalization of coherence theories and relax the requirement of orthogonality of the free states to linear independence. To be precise, we construct a resource theory in which the pure free states are a finite linearly independent set and their non-trivial superpositions are resource states. Mixed states are free if and only if they can be represented as statistical mixtures of free pure states. Thus our framework contains coherence theory as a special case. For obvious reasons, we call the free states superposition-free

and the resource states superposition states.

Such a generalization of coherence theory is interesting for several reasons. Linear independence relaxes the convenient but restrictive requirement of orthogonality, yet still provides a fundamental framework in which the notion of superposition is unambiguous and self-consistent. From a conceptual point of view, our theory helps to clarify the role of orthogonality versus linear independence. We show that many of the results of coherence theory are just special cases of their counterparts in our non-orthogonal setting. This indicates that linearly independent superposition, rather than the stronger requirement of orthogonality, is a major underlying factor in such quantum resource theories. In addition, the superposition-free states can be faithfully converted into entanglement, which implies a fundamental connection between entanglement and single-system non-classicality [12]. Thus our resource theory can give new insights into the resource theory of entanglement. Vice versa, knowledge from entanglement theory can be used to learn something about local non-classicality. As an application, the theory presented here can quantify the non-classicality in the superposition of a finite number of optical coherent states. This is not possible using the framework of coherence theory, since the optical coherent states are not orthogonal. Our theory can thus be seen as a starting point for more general resource theories with less restrictive, yet still physically meaningful constraints on the free states. Mastering these further generalizations will allow to quantify optical and other forms of non-classicality rigorously and to unify their description with entanglement theory (see also [6, 13, 17]).

This manuscript is structured as follows. In the next section, we define our free states and operations formally. Then we characterize the free operations using the concept of reciprocal states known from unambiguous state discrimination [18, 19]. This opens a possibility to prove a very convenient theorem concerning trace decreasing free operations. To ensure that they are really free, one should be able to implement the “missing” part making them trace preserving for free as well. We show that in our setting (and thus also in coherence theory), this is always possible. Afterwards we address the quantification of superposition and propose several measures. Another important result concerns the free transforma-

tions between pure states. It is shown that the maximal probability of success is almost always the solution of a semidefinite program. Finally, we investigate states with maximal superposition and the operational advantages they allow for, before concluding with a discussion on future research directions. Proofs and some additional results are given in the appendix.

Basic framework. – In this section, we give the formal definition of the free states and operations that we consider.

Definition 1. Let $\{|c_i\rangle\}_{i=1}^d$ be a normalized, linear independent and not necessarily orthogonal basis of the Hilbert space represented by \mathbb{C}^d , $d \in \mathbb{N}$. Those basis states are called *pure superposition-free states*. All density operators ρ of the form

$$\rho = \sum_{i=1}^d \rho_i |c_i\rangle \langle c_i|, \quad (1)$$

where the ρ_i form a probability distribution, are called *superposition-free*. The set of superposition-free density operators is denoted by \mathcal{F} and forms the set of *free states*. All density operators which are not superposition-free are called *superposition states* and form the set of *resource states*.

Definition 2. A Kraus operator K_n is called *superposition-free* if $K_n \rho K_n^\dagger \in \mathcal{F}$ for all $\rho \in \mathcal{F}$. Quantum operations $\Phi(\rho)$ are called *superposition-free* if they are trace preserving and can be written as

$$\Phi(\rho) = \text{tr}_B \sum_n K_n \rho_A \otimes \sigma_B K_n^\dagger, \quad (2)$$

where both σ and all K_n are superposition-free. The set of superposition-free operations forms the *free operations* and is denoted by \mathcal{FO} .

At this point, let us highlight that the definition of the free operations is not unique. This is a common trait of QRTs. The biggest possible class of free operations for our choice of the free states is given by those quantum operations that map the free states onto themselves which are denoted by \mathcal{MFO} (maximally superposition-free operations). However, these operations have the problem that, in general, they do not possess a representation in terms of superposition-free Kraus operators.

Proposition 3. \mathcal{MFO} is strictly larger than \mathcal{FO} . This is also valid in the special case of coherence theory.

Hence someone who has access to measurement outcomes of an element of \mathcal{MFO} and can thus do post-selection could conclude that a superposition-free operation generated superposition from a superposition-free state. Our definition of the free operations guarantees that one cannot create resources for free by obtaining measurement results. On the other hand, it is not as restricted as other definitions demanding for example a free dilation [20, 21]. For a discussion of alternative choices, see the appendix.

Free operations. – In order to describe the general structure of \mathcal{FO} , we need to introduce some notation. Since the pure

superposition-free states form a basis of \mathbb{C}^d , $d \in \mathbb{N}$, there exist vectors $|c_i^\perp\rangle$, $i = 1, \dots, d$ such that

$$\langle c_i^\perp | c_j \rangle = \delta_{i,j}, \quad (3)$$

which are not normalized but form a basis as well. In the context of unambiguous state discrimination, the states one gets by normalizing $|c_i^\perp\rangle$ are called *reciprocal states* [18, 19]. For explicit calculations, it is convenient to express both $\{|c_i\rangle\}_{i=1}^d$ and $\{|c_i^\perp\rangle\}_{i=1}^d$ with respect to an orthonormal basis $\{|i\rangle\}_{i=1}^d$ which will be called *computational*. Now we can introduce two linear operators V and W such that $V|i\rangle = |c_i\rangle$ and $W|i\rangle = |c_i^\perp\rangle$. Notice that both V and W are full rank since they correspond to basis transformations. From (3), it follows that $\delta_{i,j} = \langle c_i^\perp | c_j \rangle = \langle i | W^\dagger V | j \rangle$ and thus $W = (V^\dagger)^{-1}$. With this notation at hand, the explicit form of a superposition-free Kraus operator can be given, which is done in the following theorem.

Theorem 4. A Kraus operator K_n is superposition-free if and only if it is of the form

$$K_n = \sum_k c_{k,n} |c_{f_n(k)}\rangle \langle c_k^\perp|, \quad (4)$$

where $\{c_{k,n}\}$ are complex valued coefficients and $\{f_n(k)\}$ are index functions.

Incoherent Kraus operators \tilde{K}_n as defined in the limit of coherence theory [15] are thus given by $\tilde{K}_n = \sum_k c_{k,n} |f_n(k)\rangle \langle k|$ [22, 23]. If we choose the incoherent states $\{|k\rangle\}$ as the computational basis, the operator $K_n = V \tilde{K}_n V^{-1}$ has the form of a superposition-free Kraus operator. In order to have a valid, trace non-increasing quantum operation, we need

$$\mathbb{1} \geq \sum_n K_n^\dagger K_n = \sum_n (V^\dagger)^{-1} \tilde{K}_n^\dagger V^\dagger V \tilde{K}_n V^{-1}. \quad (5)$$

If the pure superposition-free states are not orthogonal, $V^\dagger \neq V^{-1}$ and in general it is therefore not possible to transform a trace non-increasing set of incoherent Kraus operators by a basis transformation V into a superposition-free one.

With the above theorem at hand, we can simplify the definition of \mathcal{FO} :

Proposition 5. All superposition-free operations $\Phi(\rho)$ can be written as $\Phi(\rho) = \sum_m F_m \rho F_m^\dagger$, where all F_m are superposition-free and no additional superposition-free system σ_B is needed.

When dealing with trace decreasing operations that can be decomposed into superposition-free Kraus operators, the question arises whether they can be seen as part of a (trace preserving) superposition-free operation. If this was not possible, it would imply that one cannot really call the trace decreasing operation free because one disregards a part that can only be done in a non-free way [24]. This leads us to our first main result.

Theorem 6. Assume we have an (incomplete) set of Kraus operator $\{K_m\}$ such that $\sum_m K_m^\dagger K_m \leq \mathbb{1}$. Then there always exist superposition-free Kraus operators $\{F_n\}$ with $\sum_m K_m^\dagger K_m + \sum_n F_n^\dagger F_n = \mathbb{1}$.

From here on we will call trace-decreasing operations with a decomposition into superposition-free Kraus operators superposition-free as well, since we can always complete them for free. Note that this is also valid in the limiting case of coherence theory.

Superposition measures. – In this section, we address the quantification of superposition, extending the method used in [15] to quantify coherence.

Definition 7. A function M mapping all quantum states to the non-negative real numbers is called a superposition measure if it is

(S1) *Faithful*

$$M(\rho) = 0 \text{ if and only if } \rho \in \mathcal{F}.$$

(S2a) *Monotonic under \mathcal{FO}*

$$M(\rho) \geq M(\Phi(\rho)) \text{ for all } \Phi \in \mathcal{FO}.$$

(S2b) *Monotonic under superposition-free selective measurements on average*

$$M(\rho) \geq \sum_n p_n M(\rho_n) :$$

$$p_n = \text{tr}(K_n \rho K_n^\dagger), \quad \rho_n = \frac{K_n \rho K_n^\dagger}{p_n}$$

$$\text{for all } \{K_n\} : \sum_n K_n^\dagger K_n = \mathbb{1}, \quad K_n \mathcal{F} K_n^\dagger \subset \mathcal{F}.$$

(S3) *Convex*

$$\sum_n p_n M(\sigma_n) \geq M\left(\sum_n p_n \sigma_n\right)$$

$$\text{for all } \{\sigma_n\}, \quad p_n \geq 0, \quad \sum_n p_n = 1.$$

If only condition (S1) and (S2a) or (S2b) are satisfied, we call M a superposition monotone.

Property (S1) demands that a state has zero superposition if and only if the state is superposition-free. As stated in (S2a), the application of a superposition-free operation to a state should not increase its superposition. If one does superposition-free selective measurements, one does not expect the superposition to increase on average which is exactly the point of (S2b). The convexity condition (S3) enforces that mixing states cannot increase the average superposition. It can be shown easily that (S2a) follows from (S2b) and (S3). As in coherence theory [15], some distance measures \mathcal{D} can be used to define superposition measures and monotones. We define a candidate $M_{\mathcal{D}}$ by

$$M_{\mathcal{D}}(\rho) = \min_{\sigma \in \mathcal{F}} \mathcal{D}(\rho, \sigma). \quad (6)$$

If \mathcal{D} is a metric, $M_{\mathcal{D}}$ fulfills (S1). If it is furthermore contractive under completely positive and trace preserving (CPTP)

maps, it fulfills (S2a) [15, 25] and for \mathcal{D} being jointly convex [26], the induced $M_{\mathcal{D}}$ fulfills condition (S3).

In accordance with [12, 14, 22], we define the superposition rank $r_S(|\psi\rangle)$ for a state $|\psi\rangle = \sum_j \psi_j |c_j\rangle$ as the number of $\psi_i \neq 0$. Assume a state $|\varphi\rangle = \sum_j \varphi_j |c_j\rangle$ can be transformed (with some probability $p > 0$) to a state $|\xi\rangle = \sum_j \xi_j |c_j\rangle$ by \mathcal{FO} . According to proposition 5, this is possible if and only if there exists a superposition-free Kraus operator $K = \sum_i c_i |c_{f(i)}\rangle \langle c_i|$ with the properties

$$\sqrt{p} \sum_i \xi_i |c_i\rangle = \sqrt{p} |\xi\rangle = K |\varphi\rangle = \sum_i \varphi_i c_i |c_{f(i)}\rangle \quad (7)$$

and $K^\dagger K \leq \mathbb{1}$. Hence the number of $\xi_i \neq 0$ is at most as large as the number of $\varphi_i \neq 0$. This proves that the superposition rank can never increase under the action of a superposition-free Kraus operator. With the definition of the superposition rank at hand, we present some explicit superposition measures.

Proposition 8. The following functions are superposition measures as defined in definition 7.

1. The relative entropy of superposition

$$M_{\text{rel.ent}}(\rho) = \min_{\sigma \in \mathcal{F}} S(\rho || \sigma), \quad (8)$$

where $S(\rho || \sigma) = \text{tr}[\rho \log \rho] - \text{tr}[\rho \log \sigma]$ denotes the quantum relative entropy. See [15] for the case of coherence theory.

2. The l_1 -measure of superposition

$$M_{l_1}(\rho) = \sum_{i \neq j} |\rho_{ij}| \quad (9)$$

for $\rho = \sum_{ij} \rho_{ij} |c_i\rangle \langle c_j|$. See again [15] for the case of coherence theory.

3. The rank-measure of superposition

$$M_{\text{rank}}(|\psi\rangle) = \log(r_S(|\psi\rangle)),$$

$$M_{\text{rank}}(\rho) = \min_{\rho = \sum_i \lambda_i |\psi_i\rangle \langle \psi_i|} \sum_i \lambda_i M_{\text{rank}}(|\psi_i\rangle). \quad (10)$$

4. The robustness of superposition

$$M_R(\rho) = \min_{\tau \text{ density matrix}} \left\{ s \geq 0 : \frac{\rho + s\tau}{1+s} \in \mathcal{F} \right\}. \quad (11)$$

This quantity has an operational interpretation in the limit of coherence theory: the robustness of coherence quantifies the advantage enabled by a quantum state in a phase discrimination task [27].

State transformations and resources. – In resource theories, it is an important question to which other states a given state can be transformed under the free operations because this leads to a hierarchy of “usefulness” in protocols. Here we consider the transformation between single copies of pure

states. Let us first clarify when probabilistic conversions are possible at all. As already mentioned, there is no possibility to increase the superposition rank of a pure state by applying a superposition-free Kraus operator. On the other hand, if two states $|\psi\rangle = \sum_{j \in R} \psi_j |c_j\rangle$ and $|\varphi\rangle = \sum_{j \in S} \varphi_j |c_j\rangle$ have the same superposition rank $r = |S| = |R|$, then there exists a superposition-free transformation that transforms one to the other with probability larger than zero. To see this, interpret R and S as (arbitrarily) ordered indexing sets. Define a function f that maps the n -th element of R to the n -th element of S and a superposition-free Kraus operator

$$K = \sqrt{p} \sum_{j \in R} \frac{\varphi_{f(j)}}{\psi_j} |c_{f(j)}\rangle \langle c_j^\perp|. \quad (12)$$

Hence $K|\psi\rangle = \sqrt{p}|\varphi\rangle$ and since $\psi_j \neq 0$ for all $j \in R$ and the pure superposition-free states $\{|c_j\rangle\}$ are linear independent, p can always be chosen such that $p > 0$ and $K^\dagger K \leq \mathbb{1}$. With the help of theorem 6, this proves that there exists a probabilistic superposition-free transformation. Different orderings of S leads to $r!$ different functions f_n and thus Kraus operators K_n . For convenience, we define

$$F_n = \sum_j \frac{\varphi_{f_n(j)}}{\psi_j} |c_{f_n(j)}\rangle \langle c_j^\perp| \quad (13)$$

with $F_n|\psi\rangle = |\phi\rangle$ and $K_n = \sqrt{p_n}F_n$. This allows us to state our second main result: The optimum free conversion probability between two pure states of the same superposition rank is the solution of the semidefinite program

$$\begin{aligned} & \text{maximize} && \sum_n p_n \\ & \text{subject to} && \sum_n p_n F_n^\dagger F_n \leq \mathbb{1} \\ & && p_n \geq 0 \quad \text{for all } n, \end{aligned} \quad (14)$$

which can be solved efficiently using numerical algorithms [28, 29]. Doing so, our investigations indicate that deterministic superposition-free transformations are rare in the case of non-orthogonal bases. Already for qubits, the probability for the existence of a deterministic transformation between two randomly picked states seems to be zero. For qubits, this is investigated analytically for a specific initial state in the appendix. If we consider superposition-free transformations to a target state with lower superposition rank than the initial state, a probabilistic transformation is still possible by the same arguments. The optimization problem however is more troublesome since we have to include Kraus operators where different pure superposition-free states are mapped to the same superposition-free target state. Therefore the optimization problem is no longer semidefinite.

If a d -dimensional superposition state can be used to generate all other d -dimensional states deterministically by means of \mathcal{FO} , it can be used for all applications. These states are said to have maximal superposition. This definition is independent of a specific superposition measure and can serve to normalize measures. Such golden units exist in coherence theory for all dimensions [15], but only for qubits in our case.

Proposition 9. *For qubit systems with $\langle c_1 | c_2 \rangle \neq 0$, there exists a single state with maximal superposition. For dimensions higher than two, there exists no state with maximal superposition in general.*

This is different to coherence theory where in dimension d , all states of the form $|m_d\rangle = 1/d \sum_{n=1}^d \exp(i\phi_n) |n\rangle$ ($\phi_n \in \mathbb{R}$) are maximally coherent [15]. A reason for this seems to be that in our more general setting, one loses entire classes of deterministic free transformations, for example diagonal unitaries which change the phases ϕ_n .

On the other hand, as in coherence theory [15], the consumption of a qubit state with maximal superposition allows to implement any unitary qubit gate by means of \mathcal{FO} .

Theorem 10. *Any unitary operation U on a qubit can be implemented by means of \mathcal{FO} and the consumption of an additional qubit state with maximal superposition $|m_2\rangle$ provided both qubits possess the same superposition-free basis. This means that for every U there exists a fixed $\Psi \in \mathcal{FO}$ independent of ρ_s acting on two qubits such that*

$$\Psi(\rho_s \otimes |m_2\rangle \langle m_2|) = (U\rho_s U^\dagger) \otimes \rho_h, \quad (15)$$

where ρ_h is a superposition-free qubit state.

This means that consuming enough qubits with maximal superposition, one can perform any unitary and thus any operation [30]

Conclusions. – We introduced a resource theory of superposition, which is a generalization of coherence theory [15]. Using the tools of quantum resource theories, we defined superposition-free states and operations. This allowed us to prove that several measures are good quantifiers of superposition, in particular the relative entropy of superposition and the l_1 -measure of superposition, which is easy to calculate. We also uncovered an important partial order structure for pure superposition states: a state can be probabilistically converted to another target state via superposition-free operations only when the target has an equal or lower superposition rank. The maximal probability for successful transformations between states of the same superposition rank is the solution of a semidefinite program. Contrasting with coherence theory, we find that only in two dimensions is there a state with maximal superposition content which can be consumed to implement an arbitrary unitary using only free operations.

Our results can help to investigate phenomena such as catalytic transformations [31–35], and act as a starting point for the investigation of mixed state transformations, transformations in the asymptotic limit [22] or approximate transformations [36]. Akin to developments in coherence theory, we can also incorporate further physical restrictions [11] such as conservation of energy [37], or restrictions for distributed scenarios such as local superposition-free operations and classical communication [38–41]. As in coherence theory [22, 39], there are also connections to entanglement theory [12] to be further understood. As potential next steps, our results could be extended to infinite dimensional states, continuous settings, or linearly dependent free states (like those found in magic state quantum computation [42, 43]). This leads towards

the ultimate goal of a fully general theory of non-classicality which puts superposition, coherence, entanglement, and quantum optical coherence on a unified standing.

We thank J. M. Matera for useful comments. This work was supported by the ERC Synergy grant BioQ, the EU project QUCHIP and an Alexander von Humboldt Professorship.

-
- [1] A. K. Ekert, *Phys. Rev. Lett.* **67**, 661 (1991).
- [2] C. H. Bennett and S. J. Wiesner, *Phys. Rev. Lett.* **69**, 2881 (1992).
- [3] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, *Phys. Rev. Lett.* **70**, 1895 (1993).
- [4] M. B. Plenio and S. Virmani, *Quant. Inf. Comp.* **7**, 1 (2007).
- [5] G. Gour and R. W. Spekkens, *New J. Phys.* **10**, 033023 (2008).
- [6] F. G. S. L. Brandão and M. B. Plenio, *Nat. Phys.* **4**, 873 (2008).
- [7] M. Horodecki and J. Oppenheim, *Int. J. Mod. Phys. B* **27**, 1345019 (2013).
- [8] J. M. Matera, D. Egloff, N. Killoran, and M. B. Plenio, *Quantum Sci. Technol.* **1**, 01LT01 (2016).
- [9] M. Hillery, *Phys. Rev. A* **93**, 012111 (2016).
- [10] G. Adesso, T. R. Bromley, and M. Cianciaruso, *J. Phys. A* **49**, 473001 (2016).
- [11] A. Streltsov, G. Adesso, and M. B. Plenio, *arXiv:1609.02439* (2016).
- [12] N. Killoran, F. E. S. Steinhoff, and M. B. Plenio, *Phys. Rev. Lett.* **116**, 080402 (2016).
- [13] W. Vogel and J. Sperling, *Phys. Rev. A* **89**, 052302 (2014).
- [14] J. Sperling and W. Vogel, *Phys. Scr.* **90**, 074024 (2015).
- [15] T. Baumgratz, M. Cramer, and M. B. Plenio, *Phys. Rev. Lett.* **113**, 140401 (2014).
- [16] J. Åberg, *arXiv:quant-ph/0612146* (2006).
- [17] K. C. Tan, T. Volkoff, H. Kwon, and H. Jeong, *arXiv:1703.01067* (2006).
- [18] A. Chefles, *Phys. Lett. A* **239**, 339 (1998).
- [19] A. Chefles, *Contemp. Phys.* **41**, 401 (2000).
- [20] E. Chitambar and G. Gour, *Phys. Rev. Lett.* **117**, 030401 (2016).
- [21] I. Marvian and R. W. Spekkens, *Phys. Rev. A* **94**, 052324 (2016).
- [22] A. Winter and D. Yang, *Phys. Rev. Lett.* **116**, 120404 (2016).
- [23] Y. Yao, X. Xiao, L. Ge, and C. P. Sun, *Phys. Rev. A* **92**, 022112 (2015).
- [24] This can happen in the case for separable operations in multipartite settings. Let for example $\{|\psi_i\rangle = |\phi_i\rangle \otimes |\xi_i\rangle\}_i$ be an unextendable separable product basis [44] of a bipartite system and define a trace-decreasing separable operation $\Lambda[\rho] = \sum_i |0\rangle\langle\phi_i| \otimes |0\rangle\langle\xi_i| \rho |\phi_i\rangle\langle 0| \otimes |\xi_i\rangle\langle 0|$ where $|0\rangle$ is an arbitrary reference state. This operation cannot be completed by separable Kraus operators by construction.
- [25] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight, *Phys. Rev. Lett.* **78**, 2275 (1997).
- [26] A. Wehrl, *Rev. Mod. Phys.* **50**, 221 (1978).
- [27] C. Napoli, T. R. Bromley, M. Cianciaruso, M. Piani, N. Johnston, and G. Adesso, *Phys. Rev. Lett.* **116**, 150502 (2016).
- [28] S. Boyd and L. Vandenberghe, *Convex optimization* (Cambridge university press, 2004).
- [29] F. Jarre and J. Stoer, *Optimierung* (Springer-Verlag, 2013).
- [30] We can express the unitary operation as a matrix with respect to the orthonormal basis obtained when applying the Gram-Schmidt process on the pure superposition-free states. As shown in [45], we can then decompose the unitary into unitaries U_2 acting on two-dimensional subspaces spanned by two pure free states. With the help of a qubit state with maximal superposition (with respect to the two free states spanning the two-dimensional subspace under consideration) every U_2 can be implemented.
- [31] D. Jonathan and M. B. Plenio, *Phys. Rev. Lett.* **83**, 3566 (1999).
- [32] S. Du, Z. Bai, and Y. Guo, *Phys. Rev. A* **91**, 052120 (2015).
- [33] J. Åberg, *Phys. Rev. Lett.* **113**, 150402 (2014).
- [34] C. Duarte, R. C. Drumond, and M. T. Cunha, *J. Phys. A* **49**, 145303 (2016).
- [35] K. Bu, U. Singh, and J. Wu, *Phys. Rev. A* **93**, 042326 (2016).
- [36] J. M. Renes, *J. Math. Phys.* **57**, 122202 (2016).
- [37] G. Chiribella and Y. Yang, *arXiv:1502.00259* (2015).
- [38] E. Chitambar and M.-H. Hsieh, *Phys. Rev. Lett.* **117**, 020402 (2016).
- [39] E. Chitambar, A. Streltsov, S. Rana, M. N. Bera, G. Adesso, and M. Lewenstein, *Phys. Rev. Lett.* **116**, 070402 (2016).
- [40] A. Streltsov, S. Rana, M. N. Bera, and M. Lewenstein, *Phys. Rev. X* **7**, 011024 (2017).
- [41] A. Streltsov, E. Chitambar, S. Rana, M. N. Bera, A. Winter, and M. Lewenstein, *Phys. Rev. Lett.* **116**, 240405 (2016).
- [42] V. Veitch, S. A. H. Mousavian, D. Gottesman, and J. Emerson, *New J. Phys.* **16**, 013009 (2014).
- [43] M. Howard and E. Campbell, *Phys. Rev. Lett.* **118**, 090501 (2017).
- [44] C. H. Bennett, D. P. DiVincenzo, T. Mor, P. W. Shor, J. A. Smolin, and B. M. Terhal, *Phys. Rev. Lett.* **82**, 5385 (1999).
- [45] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, *Phys. Rev. Lett.* **73**, 58 (1994).
- [46] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, *Rev. Mod. Phys.* **81**, 865 (2009).
- [47] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information* (Cambridge university press, 2010).
- [48] A. Pasiëka, D. W. Kribs, R. Laflamme, and R. Pereira, *Acta Appl. Math.* **108**, 697 (2009).
- [49] M.-D. Choi, *Linear Algebra Appl.* **10**, 285 (1975).
- [50] M. Jiang, S. Luo, and S. Fu, *Phys. Rev. A* **87**, 022310 (2013).
- [51] V. Vedral and M. B. Plenio, *Phys. Rev. A* **57**, 1619 (1998).
- [52] For $a = 0$, we have $K_1 = K_3 = 0$ and we need only two superposition-free Kraus operators (see also the Supplemental Material of [15]). For $a \neq 0$, in general we need four superposition-free Kraus operators to make the entire operation trace preserving. However, the actual transformation is still done by K_2 and K_4 . We will encounter a similar case in the proof of theorem 10.

Appendices

In these appendices, we give the proofs of the results in the main text and some further results. For readability, we use the short-cuts $s_\theta := \sin \theta$ and $c_\theta := \cos \theta$.

A. CHOICE OF THE FREE OPERATIONS

In this first appendix, we discuss alternative choices of the free operations defined in the main text and their relation to the free operations in coherence theory.

In the case of entanglement theory, the restriction to local operations and classical communication (LOCC) is very well motivated from a practical point of view [4]. The distant parties are allowed to perform arbitrary local quantum operations and exchange classical information but they are not allowed to transfer any quantum systems between the labs. Since classical bits cannot create entanglement, entanglement remains a resource that can be manipulated but not created. In addition, it is much cheaper to send classical information than quantum information because it can be amplified easily. However, this choice of the free operations is not unique. Different classes of free operations have been considered such as one way (forward or backward) classical communication, two way classical communication or the class of separable operations [46]. They all have their justification, either in a practical scenario or for their comparably simple mathematical structure which allows to find bound for protocols using LOCC.

Thus a debate about the choice of the free operations is necessary in every resource theory. Recently, this happened extensively in the case of coherence theory, especially since it seems difficult to justify restrictions by practical considerations (such as spacial separation in LOCC). In [11], nine different definitions of incoherent operations are collected and inclusion relations are given. The analogue of \mathcal{FO} is denoted by \mathcal{IO} and \mathcal{MFO} is equivalent to \mathcal{MIO} . One of the major concerns about \mathcal{IO} is that these operations do not possess a *free dilation* in general [20, 21]. Every quantum operation on a system A in a state ρ can be obtained from a Stinespring dilation [47]: An auxiliary system B in a state σ is introduced followed by a global unitary operation U on A and B . After a projective measurement by projectors P_m and a classical processing of the outcome, system B is discarded. According to [20], an operation possesses a free dilation if it can be obtained via a Stinespring dilation where σ and U are free and the projective measurement is a complete set of projectors on the free states. In coherence theory, the set of operations with free dilation is denoted by \mathcal{PIO} (physically incoherent operations) and has been introduced in [20]. They also showed that \mathcal{IO} is strictly larger than \mathcal{PIO} .

Whilst \mathcal{PIO} has a strong physical motivation, its power is severely reduced in comparison to \mathcal{IO} . Even the asymptotic conversion rate of the maximally coherent qubit state to any other coherent qubit state is strictly zero [20]. The generalization of \mathcal{PIO} to our framework is even more restricted. If the pure free states are not orthogonal, no complete set of projectors on the free states exists. In addition, the set of free unitary

operations is further limited as can be seen at the example of unitary operations on qubits. Unitary operations on the Bloch sphere are represented by rotations about a given axis through the origin. If the two pure free states are orthogonal and represented by $(0, 0, -1)$ and $(0, 0, 1)$, a free unitary can be decomposed into an arbitrary rotation around the z -axis and a NOT gate. If the pure free states are not orthogonal, only the equivalent to the NOT gate remains. Thus this set of free operations seems too restricted in our case to give rise to an interesting resource theory.

B. FREE OPERATIONS ON QUBITS

Geometrical interpretation of quantum operations on the Bloch sphere – For some of the proofs of the results in the main text, we make use of the geometrical interpretation of quantum operations on the Bloch sphere presented in [48]. Therefore we give a short review on this topic. Every qubit state ρ can be expanded into the Pauli basis

$$\rho = \frac{1}{2} \begin{pmatrix} 1 \\ \mathbf{r} \end{pmatrix} \sigma \quad (16)$$

with $\sigma = (\mathbb{1}, \sigma_x, \sigma_y, \sigma_z)^t$. Here σ_i denotes the Pauli matrices and $\mathbf{r} : |\mathbf{r}| \leq 1$ is a 3-component real column vector. In addition, every matrix of this form describes a valid qubit state. Every quantum operation Ψ (a linear, completely positive and trace preserving map) on the qubit can be expressed as a matrix acting on the vector of expansion coefficients. This matrix representation of Ψ is then necessarily of the form

$$\Psi = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \mathbf{t} & T \end{pmatrix}, \quad (17)$$

where \mathbf{t} is a 3-component real column vector and T is a 3×3 real matrix. However, not every operation of this form has to be a quantum operation.

The qubit operations can be decomposed into the following four geometric operations on the Bloch sphere:

1. Rotation \tilde{W}^t
2. Compression along x-, y- and z-axis to an ellipsoid with possible reflection through the y-z plane D
3. Rotation W
4. Translation \mathbf{t}

with the effect

$$\mathbf{r} \rightarrow T\mathbf{r} + \mathbf{t} = W\tilde{W}^t\mathbf{r} + \mathbf{t}. \quad (18)$$

If those operations map the Bloch sphere into itself, Ψ is positive semidefinite but not necessarily completely positive.

Superposition for qubits – Considering qubits, one can always choose a computational basis such that the superposition-free states are given by the Bloch vectors $\mathbf{r}_c = (a, 0, c)^t$ with $0 \leq a < 1$ fixed and

$$|c| \leq \sqrt{1 - a^2}. \quad (19)$$

We will use this computational basis for some of the proofs in the remainder of this Supplemental Material. The pure superposition-free states $|c_1\rangle, |c_2\rangle$ are then given by the Bloch vectors

$$\mathbf{r}_{c_1} = \begin{pmatrix} a \\ 0 \\ \sqrt{1-a^2} \end{pmatrix} \quad \text{and} \quad \mathbf{r}_{c_2} = \begin{pmatrix} a \\ 0 \\ -\sqrt{1-a^2} \end{pmatrix}. \quad (20)$$

This is equivalent to

$$\begin{aligned} |c_1\rangle &= \frac{1}{2} \left(\frac{\sqrt{1+a} + \sqrt{1-a}}{\sqrt{1+a} - \sqrt{1-a}} \right), \\ |c_2\rangle &= \frac{1}{2} \left(\frac{\sqrt{1+a} - \sqrt{1-a}}{\sqrt{1+a} + \sqrt{1-a}} \right). \end{aligned} \quad (21)$$

Since $\langle c_1 | c_2 \rangle = a$, a is a measure of the overlap of the two pure superposition-free states. To prove a difference between \mathcal{FO} and \mathcal{MFO} , we will use a certain quantum operation Φ with a matrix representation in the geometrical picture. This matrix will be defined here and in the following lemma it will be shown that this is indeed a quantum operation.

Definition 11. The matrix $\Phi = \Phi(a, \theta, \phi)$ is defined by

$$\begin{aligned} \Phi &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ \mathbf{t} & \mathbf{w} & \mathbf{0} & \mathbf{0} \end{pmatrix}, \\ \mathbf{w} &= \frac{1}{1+a} \begin{pmatrix} a - c_\phi s_\theta \\ -s_\phi s_\theta \\ -\frac{c_\theta}{2}(1+a) \end{pmatrix}, \\ \mathbf{t} &= \frac{a}{1+a} \begin{pmatrix} 1 + c_\phi s_\theta \\ s_\phi s_\theta \\ \frac{c_\theta}{2a}(1+a) \end{pmatrix}. \end{aligned} \quad (22)$$

Lemma 12. The matrix Φ represents a completely positive and trace preserving map in the geometrical picture. With the superposition-free states as defined above, it maps superposition-free states to superposition-free states.

Proof: Since the Pauli matrices are traceless, Φ is trace preserving. To show that Φ is completely positive, we will use the Choi-Jamiołkowski isomorphism [49, 50] which states that

$$C_\Psi := \sum_{i,j} |i\rangle\langle j| \otimes \Psi(|i\rangle\langle j|) \geq 0 \Leftrightarrow \Psi \text{ completely positive.} \quad (23)$$

We have

$$\begin{aligned} |1\rangle\langle 1| &= \frac{1}{2} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \sigma, & |2\rangle\langle 1| &= \frac{1}{2} \begin{pmatrix} 0 \\ 1 \\ -i \\ 0 \end{pmatrix} \sigma, \\ |2\rangle\langle 2| &= \frac{1}{2} \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix} \sigma, & |1\rangle\langle 2| &= \frac{1}{2} \begin{pmatrix} 0 \\ 1 \\ i \\ 0 \end{pmatrix} \sigma \end{aligned} \quad (24)$$

and thus

$$\Phi|1\rangle\langle 1| = \frac{1}{2} \begin{pmatrix} 1 \\ \mathbf{t} \end{pmatrix} \sigma, \quad \Phi|2\rangle\langle 1| = \frac{1}{2} \begin{pmatrix} 0 \\ \mathbf{w} \end{pmatrix} \sigma,$$

$$\Phi|2\rangle\langle 2| = \frac{1}{2} \begin{pmatrix} 1 \\ \mathbf{t} \end{pmatrix} \sigma, \quad \Phi|1\rangle\langle 2| = \frac{1}{2} \begin{pmatrix} 0 \\ \mathbf{w} \end{pmatrix} \sigma. \quad (25)$$

This allows to calculate C_Φ which is given by

$$\frac{1}{2} \begin{pmatrix} \frac{1}{2}(2+c_\theta) & \frac{a+ae^{-i\phi}s_\theta}{1+a} & -\frac{c_\theta}{2} & \frac{a-e^{-i\phi}s_\theta}{1+a} \\ \frac{a+ae^{i\phi}s_\theta}{1+a} & \frac{1}{2}(2-c_\theta) & \frac{a-e^{i\phi}s_\theta}{1+a} & \frac{c_\theta}{2} \\ -\frac{c_\theta}{2} & \frac{a-e^{-i\phi}s_\theta}{1+a} & \frac{1}{2}(2+c_\theta) & \frac{a+ae^{-i\phi}s_\theta}{1+a} \\ \frac{a-e^{i\phi}s_\theta}{1+a} & \frac{c_\theta}{2} & \frac{a+ae^{i\phi}s_\theta}{1+a} & \frac{1}{2}(2-c_\theta) \end{pmatrix}. \quad (26)$$

The eigenvalues of C_Φ are 0, 1 and $\frac{2+2a \pm R}{4(1+a)}$ with R given by

$$\begin{aligned} &\sqrt{2}\sqrt{1-2a+9a^2-(-1+a)^2c_{2\theta}+8(-1+a)ac_\theta s_\theta} \\ &\leq \sqrt{2}\sqrt{1-2a+9a^2+(1-a)^2+8(1-a)a} \\ &= 2(1-a). \end{aligned} \quad (27)$$

Using $2+2a-2(1-a) \geq 0$, all eigenvalues are larger or equal zero and thus Φ is completely positive. As a last step, it is easy to check that superposition-free states are mapped to superposition-free states since

$$\Phi \begin{pmatrix} 1 \\ a \\ 0 \\ c \end{pmatrix} = \begin{pmatrix} 1 \\ \mathbf{t} + a\mathbf{w} \end{pmatrix} = \begin{pmatrix} 1 \\ a \\ 0 \\ \frac{c_\theta}{2}(1-a) \end{pmatrix}. \quad (28)$$

■

Lemma 13. If the superposition-free states are chosen as above, the state $|m_2\rangle$ corresponding to the Bloch vector $\mathbf{r}_m = (-1, 0, 0)^t$ is for $a \neq 0$ the only candidate to have maximal superposition. The operation Φ defined in definition 11 can be used to generate all other qubit states deterministically from $|m_2\rangle$.

Proof: First we will only consider the generation of pure states from $|m_2\rangle$. The states we want to generate will be called target states. Since all pure qubit states are represented by a unit length Bloch vector, their Bloch vectors can be parametrized in polar coordinates,

$$\mathbf{r}_t = \begin{pmatrix} c_\phi s_\theta \\ s_\phi s_\theta \\ c_\theta \end{pmatrix}. \quad (29)$$

In fact it is now easy to check that

$$\Phi \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ \mathbf{t} - \mathbf{w} \end{pmatrix} = \begin{pmatrix} 1 \\ c_\phi s_\theta \\ s_\phi s_\theta \\ c_\theta \end{pmatrix} \quad (30)$$

and thus Φ transforms $\mathbf{r}_m = (-1, 0, 0)$ to the desired target state.

The generation of mixed target states ρ_M is also possible due to linearity. Since ρ_M can be decomposed into pure states through $\rho_M = \sum_i p_i |\phi_i\rangle\langle \phi_i|$, we can just apply the operation $\Phi_M = \sum_i p_i \Phi_i$ to $|m_2\rangle$ where Φ_i generates $|\phi_i\rangle$ from $|m_2\rangle$.

Finally we need to show that $|m_2\rangle$ is the only candidate to have maximal superposition. This can be seen using again the geometrical interpretation of quantum operations on the Bloch sphere. The euclidean distance between a quantum state and the set of superposition-free states is never smaller than the euclidean distance between their images under any quantum operation. The rotations and the translation preserve the distance, the compression can only reduce it.

In the case of \mathcal{MFO} , the image of the superposition-free states has to be a subset of the superposition-free states. Thus the euclidean distance between a quantum state and the set of superposition-free states cannot increase under the action of \mathcal{MFO} . Since the euclidean distance between $|m_2\rangle$ and the superposition-free states is for $a \neq 0$ larger than the euclidean distance between any other state and the superposition-free states, $|m_2\rangle$ cannot be generated with certainty from any other state by means of \mathcal{MFO} and thus not by means of \mathcal{FO} . ■

Superposition-free Kraus operators for qubits – Here we will use the results from the main text to find sufficient and necessary conditions for deterministic superposition-free operations on qubits. Remember that a superposition-free Kraus operator can be derived from an incoherent one via the transformation matrix V . Further remember that incoherent Kraus operators have at most one non-zero entry per column. Thus for qubits, there are four different types of incoherent Kraus

operators given by

$$\begin{aligned}\tilde{K}_1 &= \begin{pmatrix} \alpha & \beta \\ 0 & 0 \end{pmatrix}, \tilde{K}_2 = \begin{pmatrix} \gamma & 0 \\ 0 & \delta \end{pmatrix}, \\ \tilde{K}_3 &= \begin{pmatrix} 0 & 0 \\ \mu & \nu \end{pmatrix}, \tilde{K}_4 = \begin{pmatrix} 0 & \xi \\ \epsilon & 0 \end{pmatrix}\end{aligned}\quad (31)$$

with $\alpha, \beta, \gamma, \delta, \mu, \nu$ complex numbers. Since for Kraus operators an overall phase can be neglected, it is possible to choose in every Kraus operator one of the two non-zero entries to be real. If one chooses the computational basis $\{|1\rangle, |2\rangle\}$ in a way that

$$\begin{aligned}|c_1\rangle &= |1\rangle, \\ |c_2\rangle &= s_\theta |1\rangle + c_\theta |2\rangle, \\ 0 &\leq \theta < \frac{\pi}{2},\end{aligned}\quad (32)$$

the transformation matrix V is given by

$$V = \begin{pmatrix} 1 & s_\theta \\ 0 & c_\theta \end{pmatrix}\quad (33)$$

and

$$V^{-1} = \begin{pmatrix} 1 & -s_\theta/c_\theta \\ 0 & 1/c_\theta \end{pmatrix}, \quad V^\dagger V = \begin{pmatrix} 1 & s_\theta \\ s_\theta & 1 \end{pmatrix}.\quad (34)$$

First consider the case of a deterministic superposition-free operation in which every type of superposition-free Kraus operator $K_n = V \tilde{K}_n V^{-1}$ occurs only once. This results in the condition

$$\begin{aligned}\mathbb{1} &\stackrel{!}{=} \sum_n K_n^\dagger K_n = \sum_n (V^{-1})^\dagger \tilde{K}_n^\dagger V^\dagger V \tilde{K}_n V^{-1} \\ &= \begin{pmatrix} |\alpha|^2 + |\gamma|^2 + |\mu|^2 + |\epsilon|^2 & -s_\theta/c_\theta (|\alpha|^2 + |\gamma|^2 + |\mu|^2 + |\epsilon|^2 - \gamma^*\delta - \epsilon^*\xi) \\ & + 1/c_\theta (\alpha^*\beta + \mu^*\nu) \\ c.c. & s_\theta^2/c_\theta^2 (|\alpha|^2 + |\gamma|^2 + |\mu|^2 + |\epsilon|^2 - \gamma^*\delta - \gamma\delta^* - \epsilon^*\xi - \epsilon\xi^*) \\ & - s_\theta/c_\theta^2 (\alpha^*\beta + \alpha\beta^* + \mu^*\nu + \mu\nu^*) + 1/c_\theta^2 (|\beta|^2 + |\delta|^2 + |\nu|^2 + |\xi|^2) \end{pmatrix}\end{aligned}\quad (35)$$

where *c.c.* means the complex conjugate of the upper right matrix entry. A straight forward simplification leads to the three equations

$$\begin{aligned}1 &\stackrel{!}{=} |\alpha|^2 + |\gamma|^2 + |\mu|^2 + |\epsilon|^2, \\ 1 &\stackrel{!}{=} |\beta|^2 + |\delta|^2 + |\nu|^2 + |\xi|^2, \\ 0 &\stackrel{!}{=} \alpha^*\beta + \mu^*\nu + s_\theta (\gamma^*\delta + \epsilon^*\xi - 1).\end{aligned}\quad (36)$$

Since these equations contain θ , they seem to depend on the explicit choice of the computational basis. Now assume we had chosen another computational basis. Then it can be transformed by a unitary into the one we considered and (neglecting a physically unimportant phase) s_θ is given by $|\langle c_2|c_1\rangle|$.

Thus in general we have

$$\begin{aligned}1 &\stackrel{!}{=} |\alpha|^2 + |\gamma|^2 + |\mu|^2 + |\epsilon|^2, \\ 1 &\stackrel{!}{=} |\beta|^2 + |\delta|^2 + |\nu|^2 + |\xi|^2, \\ 0 &\stackrel{!}{=} \alpha^*\beta + \mu^*\nu + |\langle c_2|c_1\rangle| (\gamma^*\delta + \epsilon^*\xi - 1).\end{aligned}\quad (37)$$

Until now we only considered operations containing one Kraus operator of each type. In a more general scenario we can consider multiple Kraus operators of the same type and denote them by

$$\tilde{K}_{1,i} = \begin{pmatrix} \alpha_i & \beta_i \\ 0 & 0 \end{pmatrix}, \quad \tilde{K}_{2,j} = \begin{pmatrix} \gamma_j & 0 \\ 0 & \delta_j \end{pmatrix},$$

$$\tilde{K}_{3,k} = \begin{pmatrix} 0 & 0 \\ \mu_k & \nu_k \end{pmatrix}, \tilde{K}_{4,l} = \begin{pmatrix} 0 & \xi_l \\ \epsilon_l & 0 \end{pmatrix}. \quad (38)$$

Then the above equations are modified by linearity to

$$\begin{aligned} 1 &\stackrel{!}{=} \sum_i |\alpha_i|^2 + \sum_j |\gamma_j|^2 + \sum_k |\mu_k|^2 + \sum_l |\epsilon_l|^2, \\ 1 &\stackrel{!}{=} \sum_i |\beta_i|^2 + \sum_j |\delta_j|^2 + \sum_k |\nu_k|^2 + \sum_l |\xi_l|^2, \\ 0 &\stackrel{!}{=} \sum_i \alpha_i^* \beta_i + \sum_k \mu_k^* \nu_k \\ &\quad + |\langle c_2 | c_1 \rangle| \left(\sum_j \gamma_j^* \delta_j + \sum_l \epsilon_l^* \xi_l - 1 \right). \end{aligned} \quad (39)$$

That it can be useful to consider more than one superposition-free Kraus operator of the same type can be seen at the example of a quantum operation with decomposition into the two Kraus operators

$$\tilde{K}_{1,1} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \tilde{K}_{1,2} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}. \quad (40)$$

Notice that this operation is trace preserving, incoherent and uses two Kraus operators of the first type.

C. PROOFS

Here we provide the proofs of the results in the main text.

Proof of Proposition 3: In this proof we make use of the explicit representation of the superposition-free states introduced in the appendix above and the operation Φ defined in definition 11. Lemma 13 states that Φ maps superposition-free states to superposition-free states. Now it will be shown that Φ cannot be decomposed into superposition-free Kraus operators. Assume there would be such a decomposition. From theorem 4, we know how superposition-free Kraus operators are obtained from incoherent ones. In equations (31), the four different types of incoherent Kraus operators are given. With the help of equation (21) allowing to construct the matrix V , we can obtain the following four types of superposition-free qubit Kraus operators

$$\begin{aligned} K_1 &= C \begin{pmatrix} A\alpha - a\beta & -a\alpha + A\beta \\ a\alpha - B\beta & -B\alpha + a\beta \end{pmatrix}, \\ K_2 &= C \begin{pmatrix} A\gamma - B\delta & a(-\gamma + \delta) \\ -a(-\gamma + \delta) & -B\gamma + A\delta \end{pmatrix}, \\ K_3 &= C \begin{pmatrix} a\mu - B\nu & -B\mu + a\nu \\ A\mu - a\nu & -a\mu + A\nu \end{pmatrix}, \\ K_4 &= C \begin{pmatrix} a(\epsilon - \xi) & A\xi - B\epsilon \\ -B\xi + A\epsilon & -a(\epsilon - \xi) \end{pmatrix} \end{aligned} \quad (41)$$

with

$$A = 1 + \sqrt{1 - a^2}, \quad B = 1 - \sqrt{1 - a^2},$$

$$C = \frac{1}{2\sqrt{1 - a^2}}. \quad (42)$$

From equation (21), we find

$$\begin{aligned} |c_1\rangle \langle c_1| &= \frac{1}{2} \begin{pmatrix} 1 + \sqrt{1 - a^2} & a \\ a & 1 - \sqrt{1 - a^2} \end{pmatrix}, \\ |c_2\rangle \langle c_2| &= \frac{1}{2} \begin{pmatrix} 1 - \sqrt{1 - a^2} & a \\ a & 1 + \sqrt{1 - a^2} \end{pmatrix}. \end{aligned} \quad (43)$$

All superposition-free states are mapped to the same superposition-free state

$$\begin{aligned} \rho_c &= \frac{1}{2} \begin{pmatrix} 1 + \frac{c_\theta}{2}(1 - a) & a \\ a & 1 - \frac{c_\theta}{2}(1 - a) \end{pmatrix} \\ &= p |c_1\rangle \langle c_1| + q |c_2\rangle \langle c_2| \end{aligned} \quad (44)$$

with

$$\begin{aligned} p &= \frac{1}{2} + \frac{c_\theta}{4} \sqrt{\frac{1 - a}{1 + a}}, \\ q &= \frac{1}{2} - \frac{c_\theta}{4} \sqrt{\frac{1 - a}{1 + a}}. \end{aligned} \quad (45)$$

First only one Kraus operator of each type will be taken into consideration. Using

$$\begin{aligned} K_1 |c_1\rangle &= \alpha |c_1\rangle, & K_1 |c_2\rangle &= \beta |c_1\rangle, \\ K_2 |c_1\rangle &= \gamma |c_1\rangle, & K_2 |c_2\rangle &= \delta |c_2\rangle, \\ K_3 |c_1\rangle &= \mu |c_2\rangle, & K_3 |c_2\rangle &= \nu |c_2\rangle, \\ K_4 |c_1\rangle &= \epsilon |c_2\rangle, & K_4 |c_2\rangle &= \xi |c_1\rangle \end{aligned} \quad (46)$$

and

$$\sum_n K_n |c_i\rangle \langle c_i| K_n^\dagger = \rho_c \quad (47)$$

leads to

$$\begin{aligned} |\alpha|^2 + |\gamma|^2 &= p, & |\mu|^2 + |\epsilon|^2 &= q, \\ |\beta|^2 + |\xi|^2 &= p, & |\nu|^2 + |\delta|^2 &= q. \end{aligned} \quad (48)$$

Using equation (25), we can obtain the additional constraints

$$\begin{aligned} \frac{1}{2} \left(1 + \frac{c_\theta}{2} \right) &\stackrel{!}{=} \langle 1 | \left(\sum_n K_n |1\rangle \langle 1| K_n^\dagger \right) |1\rangle \\ &= \frac{1}{4(1 - a^2)} (|A\alpha - a\beta|^2 + |A\gamma - B\delta|^2 \\ &\quad + a^2|\epsilon - \xi|^2 + |a\mu - B\nu|^2), \\ \frac{1}{2} \left(1 - \frac{c_\theta}{2} \right) &\stackrel{!}{=} \langle 2 | \left(\sum_n K_n |1\rangle \langle 1| K_n^\dagger \right) |2\rangle \\ &= \frac{1}{4(1 - a^2)} (|a\alpha - B\beta|^2 + |a(\gamma - \delta)|^2 + \\ &\quad |A\epsilon - B\xi|^2 + |A\mu - a\nu|^2). \end{aligned} \quad (49)$$

With the help of equations (37) and (48) they can be simplified to

$$\begin{aligned}
& 2(1-a^2) \left(1 + \frac{c_\theta}{2}\right) \\
& \stackrel{!}{=} 2 + c_\theta(1-a) - 2a^2B - 2a(\alpha^*\beta + \alpha\beta^*)\sqrt{1-a^2} \\
& \quad - (\gamma^*\delta + \gamma\delta^* + \xi^*\epsilon + \xi\epsilon^*)a^2\sqrt{1-a^2}, \\
& 2(1-a^2) \left(1 - \frac{c_\theta}{2}\right) \\
& \stackrel{!}{=} 2 + c_\theta(1-a) - 2a^2A + 2a(\alpha^*\beta + \alpha\beta^*)\sqrt{1-a^2} \\
& \quad + (\gamma^*\delta + \gamma\delta^* + \xi^*\epsilon + \xi\epsilon^*)a^2\sqrt{1-a^2}. \quad (50)
\end{aligned}$$

Finally, adding these two equations leads to

$$c_\theta(1-a) = 0 \quad \forall \theta \quad (51)$$

which is for fixed $0 \leq a < 1$ and $\theta \neq \frac{\pi}{2}$ a contradiction to the assumption. Now consider the case that different superposition-free Kraus operators of the same type are used. Thus we use the Kraus operators introduced in equation (38). By linearity, for example $\alpha^*\beta$ will just be replaced by $\sum_i \alpha_i^* \beta_i$ in the equations above. In the step where the summation is done, the sums of coefficients will cancel out and the same contradiction is obtained. By definition, all superposition-free quantum operations map superposition-free states to superposition-free states. Thus we have proven that the set of quantum operations mapping superposition-free states to superposition-free states is greater than the set of superposition-free quantum operations. This is even true in the case of qubits. In the limit of $a = 0$, coherence theory is recovered and thus the same result holds there. ■

Proof of Theorem 4: Since both $\{|c_i\rangle\}_{i=1}^d$ and $\{|c_i^\perp\rangle\}_{i=1}^d$ form a basis, every Kraus operator L can be expanded as

$$L = \sum_{ij} L_{ij} |c_i\rangle \langle c_j^\perp|. \quad (52)$$

Applying such a general Kraus operator to an arbitrary superposition-free state $|c_k\rangle$ results in

$$L|c_k\rangle = \sum_i L_{ik} |c_i\rangle. \quad (53)$$

If L is superposition-free, L_{ik} can be non-zero for at most one i by definition. Thus L , in order to be free, has to be of the form

$$L = \sum_k c_k |c_{f(k)}\rangle \langle c_k^\perp|, \quad (54)$$

where $\{c_k\}$ are complex valued coefficients and $f(k)$ is an index function. If we have a Kraus operator of this form and apply it to a superposition-free state $\rho = \sum_i \rho_i |c_i\rangle \langle c_i|$, we find

$$\begin{aligned}
L\rho L^\dagger &= \sum_k c_k |c_{f(k)}\rangle \langle c_k^\perp| \rho |c_k\rangle \langle c_k| c_k^* |c_k^\perp\rangle \langle c_{f(k)}| \\
&= \sum_i |c_i|^2 \rho_i |c_{f(i)}\rangle \langle c_{f(i)}| \quad (55)
\end{aligned}$$

which is again superposition-free. ■

Proof of Proposition 5: In order to prove the proposition, let us consider

$$\text{tr}_B L\rho_A \otimes \sigma_B L^\dagger, \quad (56)$$

where both L and σ are superposition-free. Thus

$$\begin{aligned}
L &= \sum_{i,j} c_{i,j} |c_{g(i,j)}\rangle \langle c_{h(i,j)}^\perp|, \\
\sigma_B &= \sum_m \sigma_m |c_m\rangle \langle c_m|_B \quad (57)
\end{aligned}$$

according to theorem 4. Let $\{|x\rangle_B\}$ be an orthonormal basis of system B . Since both the pure superposition-free and the reciprocal states form a basis, we can expand

$$\begin{aligned}
\rho_A &= \sum_{s,t} \rho_{s,t} |c_s\rangle_A \langle c_t|, \\
|x\rangle_B &= \sum_l d_{x,l} |c_l^\perp\rangle_B. \quad (58)
\end{aligned}$$

Thus we find

$$\begin{aligned}
\text{tr}_B L\rho_A \otimes \sigma_B L^\dagger &= \sum_{i,j,k} c_{i,j} c_{k,j}^* \rho_{i,k} \sigma_j |c_{g(i,j)}\rangle_A \langle c_{g(k,j)}| \sum_x \langle x|c_{h(i,j)}\rangle_B \langle c_{h(k,j)}|x\rangle \\
&= \sum_{i,j,k,x,l,m} c_{i,j} c_{k,j}^* \rho_{i,k} \sigma_j |c_{g(i,j)}\rangle_A \langle c_{g(k,j)}| \delta_{l,h(i,j)} \delta_{m,h(k,j)} d_{x,l}^* d_{x,m}. \quad (59)
\end{aligned}$$

For $n = (j, x)$, we introduce superposition-free Kraus operators

$$F_n = F_{(j,x)} = \sum_i k_{i,n} |c_{f_n(i)}\rangle \langle c_i^\perp| \quad (60)$$

with

$$k_{i,n} = k_{i,(j,x)} = c_{i,j} \sum_l \delta_{l,h(i,j)} d_{x,l}^*,$$

$$f_n(i) = f_{(j,x)}(i) = g(i, j). \quad (61)$$

These Kraus operators have the property

$$\sum_{j,x} F_{j,x} \rho_A F_{j,x}^\dagger = \text{tr}_B L \rho_A \otimes \sigma_B L^\dagger \quad (62)$$

which finishes the proof by linearity. \blacksquare

Proof of Theorem 6: Let the assumptions hold. Then there exists an orthonormal basis $\{|n\rangle\}_{n=1}^d$ with

$$\sum_m K_m^\dagger K_m = \sum_{n=1}^d (1 - p_n) |n\rangle \langle n| : 0 \leq p_n \leq 1. \quad (63)$$

Remember the reciprocal vectors $|c_i^\perp\rangle$ which form a basis as well. So we can expand $|n\rangle$ in this basis

$$|n\rangle = \sum_j d_{j,n} |c_j^\perp\rangle \quad (64)$$

and write what is missing for our operation to be trace preserving as

$$\begin{aligned} \sigma_r &= \mathbb{1} - \sum_m K_m^\dagger K_m = \sum_{n=1}^d p_n |n\rangle \langle n| \\ &= \sum_{k,l=1}^d \left(\sum_{n=1}^d p_n d_{k,n} d_{l,n}^* \right) |c_k^\perp\rangle \langle c_l^\perp|. \end{aligned} \quad (65)$$

Now define additional superposition-free Kraus operators $\{F_n\}_{n=1}^d$ by

$$F_n = \sum_{k=1}^d \sqrt{p_n} d_{k,n}^* |c_1\rangle \langle c_k^\perp| \quad (66)$$

with $p_n, d_{k,n}$ from above. Since

$$\begin{aligned} \sum_{n=1}^d F_n^\dagger F_n &= \sum_{k,l,n=1}^d \sqrt{p_n} d_{k,n} \sqrt{p_n} d_{l,n}^* |c_k^\perp\rangle \langle c_1| c_1\rangle \langle c_l^\perp| \\ &= \sum_{k,l=1}^d \left(\sum_{n=1}^d p_n d_{k,n} d_{l,n}^* \right) |c_k^\perp\rangle \langle c_l^\perp| \\ &= \sigma_r, \end{aligned} \quad (67)$$

we have $\sum_m K_m^\dagger K_m + \sum_n F_n^\dagger F_n = \mathbb{1}$. \blacksquare

Proof of Proposition 8.1. The quantum relative entropy has some useful properties derived for example in [26]. For quantum states ρ and σ ,

$$S(\rho||\sigma) \geq 0, \quad (68)$$

where equality holds if and only if $\rho = \sigma$. Thus property (S1) is proven. In addition, the relative entropy is jointly convex. Therefore, as stated in the main text, $M_{\text{rel.ent}}(\rho)$ satisfies (S3). Property (S2b) can be proved as in the Supplemental Material of [15] for coherence theory. In [51], it has

been shown that the relative entropy satisfies certain conditions (called (F1)-(F5) there). Thus we can apply their theorem 2 telling us

$$\sum_n S(L_n \rho L_n^\dagger || L_n \delta L_n^\dagger) \leq S(\rho || \delta) \quad (69)$$

for any CPTP map with Kraus operator decomposition $\{L_n\}$. With their condition (F4) stating

$$\begin{aligned} \sum_n \text{tr} [L_n \rho L_n^\dagger] S \left(\frac{L_n \rho L_n^\dagger}{\text{tr} [L_n \rho L_n^\dagger]} \middle| \middle| \frac{L_n \delta L_n^\dagger}{\text{tr} [L_n \delta L_n^\dagger]} \right) \\ \leq \sum_n S(L_n \rho L_n^\dagger || L_n \delta L_n^\dagger) \end{aligned} \quad (70)$$

we find

$$\begin{aligned} \sum_n \text{tr} [L_n \rho L_n^\dagger] S \left(\frac{L_n \rho L_n^\dagger}{\text{tr} [L_n \rho L_n^\dagger]} \middle| \middle| \frac{L_n \delta L_n^\dagger}{\text{tr} [L_n \delta L_n^\dagger]} \right) \\ \leq S(\rho || \delta) \end{aligned} \quad (71)$$

again for any CPTP map with Kraus operator decomposition $\{L_n\}$. Now assume p_n and K_n as in (S2b). For a superposition-free state δ^* minimizing the relative entropy with respect to ρ we then find

$$\begin{aligned} M_{\text{rel.ent}}(\rho) &= S(\rho || \delta^*) \\ &\geq \sum_n p_n S(\rho_n || K_n \delta^* K_n^\dagger / \text{tr} [K_n \delta^* K_n^\dagger]) \\ &\geq \sum_n p_n \min_{\sigma_n \in \mathcal{F}} S(\rho_n || \sigma_n) = \sum_n p_n M_{\text{rel.ent}}(\rho_n), \end{aligned} \quad (72)$$

where we have used that $K_n \delta^* K_n^\dagger \in \mathcal{F}$. \blacksquare

Proof of Proposition 8.2: Obviously M_{l_1} maps all quantum states to the positive real numbers and (S1) is fulfilled. To prove that M_{l_1} satisfies property (S3) is straight forward. With the notation

$$\rho_n = \sum_{kl} \rho_{nkl} |c_k\rangle \langle c_l| \quad (73)$$

we have

$$\begin{aligned} M_{l_1} \left(\sum_n p_n \rho_n \right) &= \sum_{i \neq j} \left| \left(\sum_n p_n \rho_n \right)_{ij} \right| \\ &= \sum_{i \neq j} \left| \left(\sum_{kl} \left(\sum_n p_n \rho_{nkl} \right) |c_k\rangle \langle c_l| \right)_{ij} \right| \\ &= \sum_{i \neq j} \left| \sum_n p_n \rho_{n_{ij}} \right| \\ &\leq \sum_{i \neq j} \sum_n |p_n \rho_{n_{ij}}| \end{aligned}$$

$$\begin{aligned}
&= \sum_n p_n \sum_{i \neq j} |\rho_{n_{ij}}| \\
&= \sum_n p_n M_{l_1}(\rho_n).
\end{aligned} \tag{74}$$

and

$$K_n = \sum_{kl} K_{n_{kl}} |c_k\rangle \langle c_l| \tag{76}$$

The proof of property (S2b) is a bit more involved and inspired by the proof in the Supplemental Material of [15] for the case of coherence theory. We write again

$$\rho_n = \sum_{kl} \rho_{n_{kl}} |c_k\rangle \langle c_l| \tag{75}$$

alike. With this notation at hand we start with

$$\sum_n p_n M_{l_1}(\rho_n) = \sum_n p_n \sum_{i \neq j} |\rho_{n_{ij}}| = \sum_n \sum_{i \neq j} |(K_n \rho K_n^\dagger)_{ij}|. \tag{77}$$

Next step is to write down the summands $|(K_n \rho K_n^\dagger)_{ij}|$ explicitly,

$$\begin{aligned}
|(K_n \rho K_n^\dagger)_{ij}| &= \left| \left(\sum_{klstxy} K_{n_{kl}} |c_k\rangle \langle c_l| \rho_{st} |c_s\rangle \langle c_t| K_{n_{xy}}^* |c_y\rangle \langle c_x| \right)_{ij} \right| \\
&= \left| \sum_{lsty} K_{n_{il}} K_{n_{jy}}^* \rho_{st} \langle c_l | c_s \rangle \langle c_t | c_y \rangle \right| \\
&= \left| \sum_{lsy} K_{n_{il}} K_{n_{jy}}^* \rho_{ss} \langle c_l | c_s \rangle \langle c_s | c_y \rangle + \sum_{s \neq t} \sum_{ly} K_{n_{il}} K_{n_{jy}}^* \rho_{st} \langle c_l | c_s \rangle \langle c_t | c_y \rangle \right|.
\end{aligned} \tag{78}$$

With the general representation of superposition-free states

$$\rho_{cl} = \sum_i \rho_{ii} |c_i\rangle \langle c_i| \tag{79}$$

and the fact that superposition-free Kraus operators map free states to free states, we find

$$\sum_{lsy} K_{n_{il}} K_{n_{jy}}^* \rho_{ss} \langle c_l | c_s \rangle \langle c_s | c_y \rangle = (K_n \rho_{cl} K_n^\dagger)_{ij} = \delta_{ij} (K_n \rho_{cl} K_n^\dagger)_{ii}. \tag{80}$$

Now we plug equations (78) and (80) subsequently back into equation (77),

$$\begin{aligned}
\sum_n p_n M_{l_1}(\rho_n) &\stackrel{(77)}{=} \sum_n \sum_{i \neq j} |(K_n \rho K_n^\dagger)_{ij}| \\
&\stackrel{(78)}{=} \sum_n \sum_{i \neq j} \left| \sum_{lsy} K_{n_{il}} K_{n_{jy}}^* \rho_{ss} \langle c_l | c_s \rangle \langle c_s | c_y \rangle + \sum_{s \neq t} \sum_{ly} K_{n_{il}} K_{n_{jy}}^* \rho_{st} \langle c_l | c_s \rangle \langle c_t | c_y \rangle \right| \\
&\stackrel{(80)}{=} \sum_n \sum_{i \neq j} \left| \delta_{ij} (K_n \rho_{cl} K_n^\dagger)_{ii} + \sum_{s \neq t} \sum_{ly} K_{n_{il}} K_{n_{jy}}^* \rho_{st} \langle c_l | c_s \rangle \langle c_t | c_y \rangle \right| \\
&= \sum_n \sum_{i \neq j} \left| \sum_{s \neq t} \rho_{st} \sum_{ly} K_{n_{il}} K_{n_{jy}}^* \langle c_l | c_s \rangle \langle c_t | c_y \rangle \right| \\
&\leq \sum_{s \neq t} |\rho_{st}| \sum_n \sum_{i \neq j} \left| \sum_{ly} K_{n_{il}} K_{n_{jy}}^* \langle c_l | c_s \rangle \langle c_t | c_y \rangle \right|.
\end{aligned} \tag{81}$$

Regarding the last part of this expression, we find

$$\sum_n \sum_{i \neq j} \left| \sum_{ly} K_{n_{il}} K_{n_{jy}}^* \langle c_l | c_s \rangle \langle c_t | c_y \rangle \right|$$

$$\begin{aligned}
&= \sum_n \sum_{i \neq j} \left| \left(\sum_l K_{n_{il}} \langle c_l | c_s \rangle \right) \left(\sum_y K_{n_{jy}}^* \langle c_t | c_y \rangle \right) \right| \\
&\leq \sum_n \sum_i \left| \sum_l K_{n_{il}} \langle c_l | c_s \rangle \right| \left| \sum_j \left| \sum_y K_{n_{jy}}^* \langle c_t | c_y \rangle \right| \right| \\
&= \sum_n \sum_i \left| \sum_l K_{n_{il}} \langle c_l | c_s \rangle \right| \left| \sum_j \left| \sum_y K_{n_{jy}} \langle c_y | c_t \rangle \right| \right| \\
&\leq \sqrt{\left(\sum_n \left(\sum_i \left| \sum_l K_{n_{il}} \langle c_l | c_s \rangle \right| \right)^2 \right) \left(\sum_n \left(\sum_j \left| \sum_y K_{n_{jy}} \langle c_y | c_t \rangle \right| \right)^2 \right)}, \tag{82}
\end{aligned}$$

where the Cauchy-Schwarz inequality has been used in the last line. Again we will simplify a part of this expression,

$$\begin{aligned}
\sum_n \left(\sum_i \left| \sum_l K_{n_{il}} \langle c_l | c_s \rangle \right| \right)^2 &= \sum_n \sum_{ij} \left| \sum_l K_{n_{il}} \langle c_l | c_s \rangle \right| \left| \sum_k K_{n_{jk}} \langle c_k | c_s \rangle \right| \\
&= \sum_n \sum_{ij} \left| \sum_l K_{n_{il}} \langle c_l | c_s \rangle \right| \left| \sum_k K_{n_{jk}}^* \langle c_s | c_k \rangle \right| \\
&= \sum_n \sum_{ij} \left| \sum_{lk} K_{n_{il}} K_{n_{jk}}^* \langle c_l | c_s \rangle \langle c_s | c_k \rangle \right|. \tag{83}
\end{aligned}$$

Using

$$K_n |c_s\rangle \langle c_s| K_n^\dagger = \sum_{iljk} K_{n_{il}} K_{n_{jk}}^* |c_i\rangle \langle c_l| c_s \rangle \langle c_s| c_k \rangle \langle c_j|, \tag{84}$$

we can write

$$\begin{aligned}
\left| \sum_{lk} K_{n_{il}} K_{n_{jk}}^* \langle c_l | c_s \rangle \langle c_s | c_k \rangle \right| &= \left| (K_n |c_s\rangle \langle c_s| K_n^\dagger)_{ij} \right| \\
&= \left| \left(\sum_{kl} c_{kn} |c_{f_n(k)}\rangle \langle c_k^\perp | c_s \rangle \langle c_s | c_l^\perp \rangle \langle c_{f_n(l)} | c_{ln}^* \right)_{ij} \right| \\
&= \left| (|c_{sn}|^2 |c_{f_n(s)}\rangle \langle c_{f_n(s)}|)_{ij} \right| \\
&= \delta_{i,j} \delta_{i,f_n(s)} |c_{sn}|^2, \tag{85}
\end{aligned}$$

where the explicit representation of the Kraus operators from equation (4) has been used. Writing

$$\tilde{p}_n = \text{tr}(K_n |c_s\rangle \langle c_s| K_n^\dagger) = |c_{sn}|^2, \tag{86}$$

we can interpret \tilde{p}_n as the probability of outcome n when applying the operation to the state $|c_s\rangle \langle c_s|$. Thus we have

$$1 = \sum_n \tilde{p}_n = \sum_n |c_{sn}|^2. \tag{87}$$

Putting the pieces together leads to

$$\begin{aligned}
\sum_n \left(\sum_i \left| \sum_l K_{n_{il}} \langle c_l | c_s \rangle \right| \right)^2 &\stackrel{(83)}{=} \sum_n \sum_{ij} \left| \sum_{lk} K_{n_{il}} K_{n_{jk}}^* \langle c_l | c_s \rangle \langle c_s | c_k \rangle \right| \\
&\stackrel{(85)}{=} \sum_{nij} \delta_{i,j} \delta_{i,f_n(s)} |c_{sn}|^2
\end{aligned}$$

$$= \sum_n |c_{sn}|^2 \stackrel{(87)}{=} 1. \quad (88)$$

Finally we can finish the proof through

$$\begin{aligned} \sum_n p_n M_{l_1}(\rho_n) &\stackrel{(81)}{\leq} \sum_{s \neq t} |\rho_{st}| \sum_n \sum_{i \neq j} \left| \sum_{ly} K_{n_{il}} K_{n_{jy}}^* \langle c_l | c_s \rangle \langle c_t | c_y \rangle \right| \\ &\stackrel{(82)}{\leq} \sum_{s \neq t} |\rho_{st}| \sqrt{\left(\sum_n \left(\sum_i \left| \sum_l K_{n_{il}} \langle c_l | c_s \rangle \right| \right)^2 \right) \left(\sum_n \left(\sum_j \left| \sum_y K_{n_{jy}} \langle c_y | c_t \rangle \right| \right)^2 \right)} \\ &\stackrel{(88)}{=} \sum_{s \neq t} |\rho_{st}| = M_{l_1}(\rho). \end{aligned} \quad (89)$$

■

Proof of Proposition 8.3: It can be seen directly that (S1) is fulfilled. To show (S2b), let p_n , ρ_n and K_n be as in the definition. Assume without loss of generality $p_n \neq 0 \forall n$. Define for every state ρ a decomposition into pure states by $\rho = \sum_i \tilde{\lambda}_i |\tilde{\psi}_i\rangle \langle \tilde{\psi}_i|$ such that

$$M_{\text{rank}}(\rho) = \sum_i \tilde{\lambda}_i M_{\text{rank}}(|\tilde{\psi}_i\rangle). \quad (90)$$

Further define

$$|\tilde{\psi}_i^{(n)}\rangle = \frac{K_n |\tilde{\psi}_i\rangle}{\sqrt{p_n}}. \quad (91)$$

Now notice that

$$\begin{aligned} \rho_n &= \frac{K_n \rho K_n^\dagger}{p_n} = \sum_i \frac{\tilde{\lambda}_i}{p_n} K_n |\tilde{\psi}_i\rangle \langle \tilde{\psi}_i| K_n^\dagger \\ &= \sum_i \tilde{\lambda}_i |\tilde{\psi}_i^{(n)}\rangle \langle \tilde{\psi}_i^{(n)}|. \end{aligned} \quad (92)$$

Remember that the superposition rank can never increase under the action of a superposition-free Kraus operator. Since $\log(r_S)$ decreases if r_S decreases, we can finish the proof of (S2b) by

$$\begin{aligned} M_{\text{rank}}(\rho) &= \sum_i \tilde{\lambda}_i M_{\text{rank}}(|\tilde{\psi}_i\rangle) \\ &= \sum_n p_n \sum_i \tilde{\lambda}_i M_{\text{rank}}(|\tilde{\psi}_i\rangle) \\ &\geq \sum_n p_n \sum_i \tilde{\lambda}_i M_{\text{rank}}(|\tilde{\psi}_i^{(n)}\rangle) \\ &\geq \sum_n p_n \min_i \sum_i \lambda_i^{(n)} M_{\text{rank}}(|\psi_i^{(n)}\rangle) \\ &= \sum_n p_n M_{\text{rank}}(\rho_n), \end{aligned} \quad (93)$$

where the minimization in the second last line runs over all decompositions $\rho_n = \sum_i \lambda_i^{(n)} |\psi_i^{(n)}\rangle \langle \psi_i^{(n)}|$. In order to

show (S3), follow the same spirit. Again, define $\sigma_n = \sum_i \tilde{\lambda}_i^{(n)} |\tilde{\psi}_i^{(n)}\rangle \langle \tilde{\psi}_i^{(n)}|$ such that

$$M_{\text{rank}}(\sigma_n) = \sum_i \tilde{\lambda}_i^{(n)} M_{\text{rank}}(|\tilde{\psi}_i^{(n)}\rangle) \quad (94)$$

and note that

$$\sum_n p_n \sigma_n = \sum_{n,i} p_n \tilde{\lambda}_i^{(n)} |\tilde{\psi}_i^{(n)}\rangle \langle \tilde{\psi}_i^{(n)}|. \quad (95)$$

Hence

$$\begin{aligned} M_{\text{rank}}\left(\sum_n p_n \sigma_n\right) &= \min_{\sum_n p_n \sigma_n = \sum_i \lambda_i |\psi_i\rangle \langle \psi_i|} \sum_i \lambda_i M_{\text{rank}}(|\psi_i\rangle) \\ &\leq \sum_{n,i} p_n \tilde{\lambda}_i^{(n)} M_{\text{rank}}(|\tilde{\psi}_i^{(n)}\rangle) \\ &= \sum_n p_n M_{\text{rank}}(\sigma_n), \end{aligned} \quad (96)$$

which finishes the proof. ■

Proof of Proposition 8.4: Property (S1) is given by definition. To prove (S2b) decompose

$$\rho = (1 + \tilde{s})\tilde{\delta} - \tilde{s}\tilde{\tau} \quad (97)$$

such that $\tilde{s} = M_R(\rho)$, $\tilde{\delta} \in \mathcal{F}$ and $\tilde{\tau}$ a density matrix. Hence

$$\begin{aligned} K_n \rho K_n^\dagger &= (1 + \tilde{s}) K_n \tilde{\delta} K_n^\dagger - \tilde{s} K_n \tilde{\tau} K_n^\dagger, \\ p_n &= \text{tr}(K_n \rho K_n^\dagger) \\ &= (1 + \tilde{s}) \text{tr}(K_n \tilde{\delta} K_n^\dagger) - \tilde{s} \text{tr}(K_n \tilde{\tau} K_n^\dagger). \end{aligned} \quad (98)$$

Assume without loss of generality $p_n \neq 0$, $\text{tr}(K_n \tilde{\delta} K_n^\dagger) \neq 0$ and $\text{tr}(K_n \tilde{\tau} K_n^\dagger) \neq 0$. Then

$$\rho_n = \frac{K_n \rho K_n^\dagger}{p_n}$$

$$\begin{aligned}
&= \frac{1+\tilde{s}}{p_n} \frac{\text{tr}(K_n \tilde{\delta} K_n^\dagger)}{\text{tr}(K_n \tilde{\delta} K_n^\dagger)} \frac{K_n \tilde{\delta} K_n^\dagger}{\text{tr}(K_n \tilde{\delta} K_n^\dagger)} \\
&\quad - \frac{\tilde{s}}{p_n} \frac{\text{tr}(K_n \tilde{\tau} K_n^\dagger)}{\text{tr}(K_n \tilde{\tau} K_n^\dagger)} \frac{K_n \tilde{\tau} K_n^\dagger}{\text{tr}(K_n \tilde{\tau} K_n^\dagger)} \\
&= (1+s_n) \tilde{\delta}^{(n)} - s_n \tilde{\tau}^{(n)} \quad (99)
\end{aligned}$$

for $s_n = \frac{\tilde{s} \text{tr}(K_n \tilde{\tau} K_n^\dagger)}{p_n} \geq 0$, $\tilde{\delta}^{(n)} = \frac{K_n \tilde{\delta} K_n^\dagger}{\text{tr}(K_n \tilde{\delta} K_n^\dagger)} \in \mathcal{F}$ and a density matrix $\tilde{\tau}^{(n)} = \frac{K_n \tilde{\tau} K_n^\dagger}{\text{tr}(K_n \tilde{\tau} K_n^\dagger)}$. Thus $M_R(\rho_n) \leq s_n$ and

$$\begin{aligned}
\sum_n p_n M_R(\rho_n) &\leq \sum_n p_n s_n = \sum_n \tilde{s} \text{tr}(K_n \tilde{\tau} K_n^\dagger) \\
&= \tilde{s} \text{tr} \left(\sum_n K_n \tilde{\tau} K_n^\dagger \right) = \tilde{s} = M_R(\rho). \quad (100)
\end{aligned}$$

In order to prove (S3), decompose again two density matrices ρ_1, ρ_2 into

$$\rho_i = (1 + \tilde{s}_i) \tilde{\delta}_i - \tilde{s}_i \tilde{\tau}_i \quad (101)$$

such that $\tilde{s}_i = M_R(\rho_i)$, $\tilde{\delta}_i \in \mathcal{F}$ and $\tilde{\tau}_i$ a density matrix. For $0 \leq p \leq 1$ and without loss of generality $\tilde{s}_i \neq 0$,

$$\begin{aligned}
&p\rho_1 + (1-p)\rho_2 \\
&= p \left[(1 + \tilde{s}_1) \tilde{\delta}_1 - \tilde{s}_1 \tilde{\tau}_1 \right] + (1-p) \left[(1 + \tilde{s}_2) \tilde{\delta}_2 - \tilde{s}_2 \tilde{\tau}_2 \right] \\
&= [1 + p\tilde{s}_1 + (1-p)\tilde{s}_2] \frac{p(1 + \tilde{s}_1) \tilde{\delta}_1 + (1-p)(1 + \tilde{s}_2) \tilde{\delta}_2}{1 + p\tilde{s}_1 + (1-p)\tilde{s}_2} \\
&\quad - [p\tilde{s}_1 + (1-p)\tilde{s}_2] \frac{p\tilde{s}_1 \tilde{\tau}_1 + (1-p)\tilde{s}_2 \tilde{\tau}_2}{p\tilde{s}_1 + (1-p)\tilde{s}_2} \\
&= (1+s)\delta - s\tau \quad (102)
\end{aligned}$$

for $s = p\tilde{s}_1 + (1-p)\tilde{s}_2 \geq 0$, $\delta = \frac{p(1+\tilde{s}_1)\tilde{\delta}_1 + (1-p)(1+\tilde{s}_2)\tilde{\delta}_2}{1+p\tilde{s}_1 + (1-p)\tilde{s}_2} \in \mathcal{F}$ and a density matrix $\tau = \frac{p\tilde{s}_1 \tilde{\tau}_1 + (1-p)\tilde{s}_2 \tilde{\tau}_2}{p\tilde{s}_1 + (1-p)\tilde{s}_2}$. Hence

$$\begin{aligned}
&M_R(p\rho_1 + (1-p)\rho_2) \\
&= \min \{t \geq 0 : p\rho_1 + (1-p)\rho_2 = (1+t)\delta - t\tau : \\
&\quad \delta \in \mathcal{F}, \tau \text{ density matrix}\} \\
&\leq s \\
&= p\tilde{s}_1 + (1-p)\tilde{s}_2 \\
&= pM_R(\rho_1) + (1-p)M_R(\rho_2) \quad (103)
\end{aligned}$$

which proves convexity. \blacksquare

Proof of Proposition 9: To prove this proposition, we will use again the representation introduced in the appendix B and the four different types of superposition-free Kraus operators given in equation (41). To show that $|m_2\rangle$, the candidate from lemma 13, has indeed maximal superposition, we will explicitly construct a superposition-free operation that generates a target state

$$|\psi_t\rangle = \begin{pmatrix} c_{\theta/2} \\ e^{i\phi} s_{\theta/2} \end{pmatrix} \quad (104)$$

from $|m_2\rangle$ with certainty. In order to achieve this, we choose one superposition-free Kraus operator of each type with

$$\begin{aligned}
\delta &= \frac{1}{2(1+a)} [(B+a)c_{\theta/2} - (a+A)e^{i\phi} s_{\theta/2}], \\
\gamma &= \frac{1}{2(1+a)} [(A+a)c_{\theta/2} - (a+B)e^{i\phi} s_{\theta/2}], \\
\epsilon &= -\frac{1}{2(1+a)} [(B+a)c_{\theta/2} - (a+A)e^{i\phi} s_{\theta/2}], \\
\xi &= -\frac{1}{2(1+a)} [(A+a)c_{\theta/2} - (a+B)e^{i\phi} s_{\theta/2}], \\
\alpha &= \beta = \mu = \nu = \sqrt{\frac{a}{2(1+a)}(1+c_\phi s_\theta)}, \quad (105)
\end{aligned}$$

where A and B are defined as in equation (42). With the help of equation (37), it is easy to check that they form a trace preserving operation, i.e. $\sum_{n=1}^4 K_n^\dagger K_n = \mathbb{1}$. Remember that in the representation chosen, $|m_2\rangle$ is given by

$$|m_2\rangle = \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \quad (106)$$

Thus one can see directly that the four Kraus operators have the property

$$\begin{aligned}
K_2 |m_2\rangle &= K_4 |m_2\rangle = \frac{1}{\sqrt{2}} |\psi_t\rangle, \\
K_1 |m_2\rangle &= K_3 |m_2\rangle = 0. \quad (107)
\end{aligned}$$

With the help of theorem 6, the explicit construction of K_1 and K_3 would not have been necessary [52]. The generation of mixed target states ρ_M is again possible due to linearity. Since ρ_M can be decomposed into pure states through $\rho_M = \sum_i p_i |\phi_i\rangle \langle \phi_i|$, we can just apply the operation $\Phi_M = \sum_i p_i \Phi_i$ to $|m_2\rangle$ where Φ_i generates $|\phi_i\rangle$ from $|m_2\rangle$. As stated in lemma 13, $|m_2\rangle$ is the only qubit state with maximal superposition for $\langle c_1 | c_2 \rangle \neq 0$.

To prove the second part of the proposition, we show with the help of the l_1 -measure of superposition and some tools from optimization theory (described for example in [28]) that for a specific superposition-free basis in $d=3$, there exists no state with maximal superposition.

The main idea is that a state with maximal superposition has to maximize the l_1 -measure of superposition since a superposition measure cannot increase under superposition-free operations. First step is to define the superposition-free states in a fixed dimension d . Then all states maximizing M_{l_1} have to be determined. Once these candidate states are found, one can use the optimization problem from the main text to calculate the maximal transformation probability to a given target state. If the solution is smaller than one, the considered candidate state cannot have maximal superposition. If all candidate states are ruled out, one concludes that there is no state with maximal superposition (for this set of superposition-free states). For specific choices of superposition-free states, this can be done analytically using the concept of duality. In the case of $d=3$, we choose the pure superposition-free states to

be represented by

$$|c_1\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, |c_2\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, |c_3\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}. \quad (108)$$

Then the (not normalized) reciprocal vectors introduced in the main text are given by

$$|c_1^\perp\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}, |c_2^\perp\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \\ |c_3^\perp\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}. \quad (109)$$

It is easy to check that

$$\langle c_i | c_j \rangle = \frac{1}{2} (1 + \delta_{i,j}). \quad (110)$$

Because of this symmetry, every candidate state has to be of the form

$$|m_d\rangle = \mathcal{N} \sum_i e^{i\phi_i} |c_i\rangle, \quad (111)$$

where \mathcal{N} is positive and $\phi_3 = 0$ by choice. The l_1 measure of superposition is then given by

$$M_{l_1}(|m_d\rangle) = 3(3-1)\mathcal{N}^2 = 6\mathcal{N}^2. \quad (112)$$

Thus to find proper candidate states maximizing M_{l_1} , we have to solve the following optimization problem

$$\begin{aligned} &\text{maximize} && M_{l_1}(|m_d\rangle) = 6\mathcal{N}^2 \\ &\text{subject to} && 1 = \langle m_d | m_d \rangle \\ &&& = \mathcal{N}^2 [3 + s_{\phi_1} s_{\phi_2} + c_{\phi_1} c_{\phi_2} + c_{\phi_1} + c_{\phi_2}]. \end{aligned} \quad (113)$$

This is equivalent to minimizing $s_{\phi_1} s_{\phi_2} + c_{\phi_1} c_{\phi_2} + c_{\phi_1} + c_{\phi_2}$. We choose $\phi_{1,2} \in [0, 2\pi)$ and do the change of variables

$$\alpha = \frac{\phi_1 - \phi_2}{2} \in (-\pi, \pi], \quad \beta = \frac{\phi_1 + \phi_2}{2} \in [0, 2\pi). \quad (114)$$

This transforms the initial optimization problem (113) into the problem

$$\begin{aligned} &\text{minimize} && c_{2\alpha} + 2c_\alpha c_\beta \\ &\text{subject to} && \alpha \in (-\pi, \pi] \\ &&& \beta \in [0, 2\pi), \end{aligned} \quad (115)$$

which can be solved considering three separate cases.

First case ($c_\alpha > 0$): In this case, in order to minimize $c_{2\alpha} + 2c_\alpha c_\beta$, we need $c_\beta = -1 \Leftrightarrow \beta = \pi$. The problem reduces to

$$\text{minimize} \quad c_{2\alpha} - 2c_\alpha$$

$$\text{subject to} \quad \alpha \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right). \quad (116)$$

The solution is $-\frac{3}{2}$ for $\alpha = \pm\frac{\pi}{3}$ which leads to

$$\begin{aligned} \mathcal{N}^{(1)} &= \sqrt{\frac{2}{3}}, & \phi_1^{(1)} &= \frac{4\pi}{3}, & \phi_2^{(1)} &= \frac{2\pi}{3}, \\ \mathcal{N}^{(2)} &= \sqrt{\frac{2}{3}}, & \phi_1^{(2)} &= \frac{2\pi}{3}, & \phi_2^{(2)} &= \frac{4\pi}{3}. \end{aligned} \quad (117)$$

Second case ($c_\alpha < 0$): In order to minimize $c_{2\alpha} + 2c_\alpha c_\beta$, we now need $c_\beta = 1 \Leftrightarrow \beta = 0$ and the remaining problem is

$$\begin{aligned} &\text{minimize} && c_{2\alpha} + 2c_\alpha \\ &\text{subject to} && \alpha \in \left(-\pi, -\frac{\pi}{2}\right) \cup \left(-\frac{\pi}{2}, \pi\right). \end{aligned} \quad (118)$$

Again, the solution is $-\frac{3}{2}$, this time for $\alpha = \pm\frac{2\pi}{3}$ which leads to

$$\begin{aligned} \mathcal{N}^{(3)} &= \sqrt{\frac{2}{3}}, & \phi_1^{(3)} &= \frac{2\pi}{3}, & \phi_2^{(3)} &= -\frac{2\pi}{3}, \\ \mathcal{N}^{(4)} &= \sqrt{\frac{2}{3}}, & \phi_1^{(4)} &= -\frac{2\pi}{3}, & \phi_2^{(4)} &= \frac{2\pi}{3}. \end{aligned} \quad (119)$$

Third case ($c_\alpha = 0 \Leftrightarrow \alpha = \pm\frac{\pi}{2}$): This case is trivial since then $c_{2\alpha} + 2c_\alpha c_\beta$ is equal to -1 for all β and no more global minima are found.

Thus we remain with four candidate states which maximize M_{l_1} . They are given by

$$\begin{aligned} |m_d^{(1)}\rangle &= \sqrt{\frac{2}{3}} \left[e^{i\frac{4\pi}{3}} |c_1\rangle + e^{i\frac{2\pi}{3}} |c_2\rangle + |c_3\rangle \right], \\ |m_d^{(2)}\rangle &= \sqrt{\frac{2}{3}} \left[e^{i\frac{2\pi}{3}} |c_1\rangle + e^{i\frac{4\pi}{3}} |c_2\rangle + |c_3\rangle \right], \\ |m_d^{(3)}\rangle &= \sqrt{\frac{2}{3}} \left[e^{i\frac{2\pi}{3}} |c_1\rangle + e^{-i\frac{2\pi}{3}} |c_2\rangle + |c_3\rangle \right], \\ |m_d^{(4)}\rangle &= \sqrt{\frac{2}{3}} \left[e^{-i\frac{2\pi}{3}} |c_1\rangle + e^{i\frac{2\pi}{3}} |c_2\rangle + |c_3\rangle \right]. \end{aligned} \quad (120)$$

Next step is to show that none of these states can be transformed to every other state with certainty using superposition-free Kraus operators. Therefore we will focus on the transformation to target states of the form $|\psi\rangle = \sum_{k=1}^3 \psi_k |c_k\rangle$ with $\psi_k \neq 0 \forall k$ since then we can use the semidefinite program introduced in the main text:

$$\begin{aligned} &\text{maximize} && \sum_n p_n \\ &\text{subject to} && \sum_n p_n F_n^\dagger F_n \leq \mathbb{1} \\ &&& p_n \geq 0 \quad \forall n. \end{aligned} \quad (121)$$

Note that $\{F_n\}$ depends on the choice of the target states and on the candidate state we want to rule out. This problem is equivalent to

$$\text{minimize} \quad -\sum_n p_n$$

$$\begin{aligned} \text{subject to} \quad & \sum_n p_n F_n^\dagger F_n - \mathbb{1} \leq 0 \\ & p_n \geq 0 \quad \forall n. \end{aligned} \quad (122)$$

If we interpret the condition $p_n \geq 0 \forall n$ as a restriction of the domain of the objective function, the Lagrangian is given by

$$\begin{aligned} L(p_n, \Lambda) &= -\sum_n p_n + \text{tr} \left(\Lambda \left(\sum_n p_n F_n^\dagger F_n - \mathbb{1} \right) \right) \\ &= \text{tr} \left(-\frac{1}{d} \mathbb{1} \sum_n p_n + \Lambda \left(\sum_n p_n F_n^\dagger F_n - \mathbb{1} \right) \right) \end{aligned} \quad (123)$$

and the dual function by

$$\begin{aligned} g(\Lambda) &= \inf_{p_n \geq 0} L(p_n, \Lambda) \\ &= -\text{tr}(\Lambda) + \inf_{p_n \geq 0} \sum_n p_n \text{tr} \left(-\frac{1}{d} \mathbb{1} + \Lambda F_n^\dagger F_n \right) \\ &= \begin{cases} -\infty & \text{if } \exists n : \text{tr} \left(-\frac{1}{d} \mathbb{1} + \Lambda F_n^\dagger F_n \right) < 0 \\ -\text{tr}(\Lambda) & \text{if } \text{tr} \left(-\frac{1}{d} \mathbb{1} + \Lambda F_n^\dagger F_n \right) \geq 0 \forall n \end{cases}. \end{aligned} \quad (124)$$

Hence the dual problem is

$$\begin{aligned} \text{maximize} \quad & -\text{tr}(\Lambda) \\ \text{subject to} \quad & \text{tr}(\Lambda F_n^\dagger F_n) \geq 1 \quad \forall n \\ & \Lambda \geq 0. \end{aligned} \quad (125)$$

By duality, we have

$$-\text{tr} \Lambda \leq -\sum_n p_n \quad (126)$$

or

$$\sum_n p_n \leq \text{tr} \Lambda \quad (127)$$

for all feasible Λ . Thus if one can find a feasible Λ with $\text{tr}(\Lambda) < 1$, it is shown that there exists no deterministic free transformation between the candidate and the target state. For $|\psi\rangle = \sum_i \psi_i |c_i\rangle$ we have $\psi_i = \langle c_i^\perp | \psi \rangle$. Now we choose as a target state

$$|\psi\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}} \left(-|c_1\rangle + |c_2\rangle + |c_3\rangle \right). \quad (128)$$

A direct calculation (using for example Mathematica) shows that in this case, we have $\text{tr}(F_n^\dagger F_n) = \frac{51}{16}$ for all n and for all candidate states. If we choose

$$\Lambda = \frac{t}{3} \mathbb{1}_3, \quad (129)$$

we have $\text{tr} \Lambda = t$ and $\Lambda \geq 0$ iff $t \geq 0$. Thus

$$\text{tr}(\Lambda F_n^\dagger F_n) = \frac{t}{3} \text{tr}(F_n^\dagger F_n) = t \frac{17}{16} \quad \forall n, \forall |m_d^{(i)}\rangle. \quad (130)$$

Now we can choose $t = \frac{16}{17}$ for Λ to be feasible and bound the maximal conversion probability from above by $p_{\max} \leq \text{tr} \Lambda = t = \frac{16}{17} < 1$. ■

Proof of theorem 10: In this proof, we will again make use of the computational basis introduced in the appendix B. A matrix representation of a unitary U with respect to this basis can be represented as

$$U = \begin{pmatrix} u_{00} & u_{01} \\ u_{10} & u_{11} \end{pmatrix} = e^{i\phi_g} \begin{pmatrix} e^{i\phi_1} c_\theta & e^{i\phi_2} s_\theta \\ -e^{-i\phi_2} s_\theta & e^{-i\phi_1} c_\theta \end{pmatrix}, \quad (131)$$

where ϕ_g is a (physically) unimportant global phase. Therefore we will drop it from now on. Consider the two Kraus operators

$$\begin{aligned} F_0 &= c_{00} |c_1 c_1\rangle \langle c_1^\perp c_1^\perp| + c_{10} |c_2 c_1\rangle \langle c_1^\perp c_2^\perp| \\ &\quad + c_{01} |c_1 c_1\rangle \langle c_2^\perp c_1^\perp| + c_{11} |c_2 c_1\rangle \langle c_2^\perp c_2^\perp|, \\ F_1 &= d_{00} |c_1 c_2\rangle \langle c_1^\perp c_2^\perp| + d_{10} |c_2 c_2\rangle \langle c_1^\perp c_1^\perp| \\ &\quad + d_{01} |c_1 c_2\rangle \langle c_2^\perp c_2^\perp| + d_{11} |c_2 c_2\rangle \langle c_2^\perp c_1^\perp|. \end{aligned} \quad (132)$$

We fix the coefficients by

$$\begin{aligned} c_{00} &= \frac{A u_{00} + a(u_{01} - u_{10}) - B u_{11}}{2\sqrt{1+a}}, \\ c_{01} &= \frac{A u_{01} + a(u_{00} - u_{11}) - B u_{10}}{2\sqrt{1+a}}, \\ c_{10} &= \frac{B u_{01} + a(u_{00} - u_{11}) - A u_{10}}{2\sqrt{1+a}}, \\ c_{11} &= \frac{B u_{00} + a(u_{01} - u_{10}) - A u_{11}}{2\sqrt{1+a}}, \\ d_{00} &= \frac{B u_{11} + a(u_{10} - u_{01}) - A u_{00}}{2\sqrt{1+a}}, \\ d_{01} &= \frac{B u_{10} + a(u_{11} - u_{00}) - A u_{01}}{2\sqrt{1+a}}, \\ d_{10} &= \frac{A u_{10} + a(u_{11} - u_{00}) - B u_{01}}{2\sqrt{1+a}}, \\ d_{11} &= \frac{A u_{11} + a(u_{10} - u_{01}) - B u_{00}}{2\sqrt{1+a}}. \end{aligned} \quad (133)$$

Thus the coefficients depend on the unitary transformation U and on the overlap a of the pure superposition-free states (also through A and B which are defined as in equation (42)). The two Kraus operators F_0 and F_1 are superposition-free because they satisfy theorem 4. In addition, we have for every qubit state $|s\rangle$

$$\begin{aligned} F_0 |s\rangle \otimes |m_2\rangle &= \frac{1}{\sqrt{2}} (U |s\rangle) \otimes |c_1\rangle, \\ F_1 |s\rangle \otimes |m_2\rangle &= \frac{1}{\sqrt{2}} (U |s\rangle) \otimes |c_2\rangle. \end{aligned} \quad (134)$$

Making use of the explicit representation of U , the eigenvalues of $F_0^\dagger F_0 + F_1^\dagger F_1$ can be calculated to be $(1, 1, 1, (\frac{1-a}{1+a})^2)$. With the help of theorem 6, we know

that there exist additional Kraus operators $\{L_i\}$ such that $F_0^\dagger F_0 + F_1^\dagger F_1 + \sum_i L_i^\dagger L_i = \mathbb{1}$ and $L_i |s\rangle \otimes |m_2\rangle = 0$. By linearity, this finishes the proof. ■

D. ON SUPERPOSITION TRANSFORMATIONS

In this appendix, we have a short look at superposition-free transformations of qubit states. Therefore we use again the representation introduced in the appendix B, the semidefinite program from the main text and its dual problem (125). For qubits, there are only two different useful Kraus operators contributing to the transformation except in the case where we transform to the two pure superposition-free states. Now we define the Bloch representation of the initial state $|\psi\rangle$ and the target state $|\phi\rangle$ by

$$\begin{aligned} |\psi\rangle &= c_{w/2} |1\rangle + s_{w/2} e^{iy} |2\rangle, \\ |\phi\rangle &= c_{x/2} |1\rangle + s_{x/2} e^{iz} |2\rangle. \end{aligned} \quad (135)$$

Further we consider

$$\Lambda = \frac{t}{2} \mathbb{1}_2 \quad (136)$$

in the dual problem. Thus $\text{tr} \Lambda = t$ and $\Lambda \geq 0$ iff $t \geq 0$. For $a = 1/2$, $w = \pi/2$, $y = 0$, one finds

$$\text{tr} (\Lambda F_1^\dagger F_1) = \text{tr} (\Lambda F_2^\dagger F_2) = t [3 - 2c_z s_x] \geq t. \quad (137)$$

Remember $x \in [0, \pi]$ and $z \in [0, 2\pi]$ due to the definition of the Bloch representation. Thus we only have equality for $x = \pi/2$ and $z = 0$ which is equivalent to $|\phi\rangle = |\psi\rangle$. In case the above expression is strictly larger than t , we can always choose $t < 1$ such that $\text{tr} (\Lambda F_1^\dagger F_1) = 1$. Then Λ is feasible and the optimal probability of successfully transforming the initial state to the target state is smaller than one. For transformation to the pure superposition-free states, we have to consider additional Kraus operators. Thus finally we can conclude that there are only the three trivial pure states to which $|\psi\rangle = \frac{1}{\sqrt{2}}(|1\rangle + |2\rangle) = \frac{1}{\sqrt{3}}(|c_1\rangle + |c_2\rangle)$ can be transformed by \mathcal{FO} with certainty: to itself and to the pure superposition-free states.

This is surprising because as one can see easily with the help of the l_1 -measure of superposition, the initial state under consideration contains a large amount of superposition and there are other states with less superposition to which this state cannot be transformed by \mathcal{FO} with certainty nevertheless. In contrast, for $a = 0$, it can be shown with the help of [32] that a pure state can be transformed deterministically to all other pure states that are closer or equally close to the z -axis of the Bloch sphere (and have thus less superposition according to the l_1 -measure of superposition). A possible explanation to this difference could be that by breaking the symmetry on the Bloch sphere, one loses an entire class of superposition-free operations since rotations around the z -axis are no longer possible.